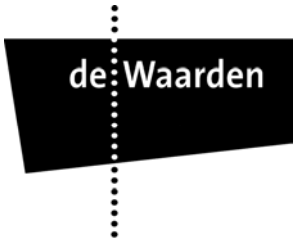
THE PERCEPTION OF CHILD PROBLEM BEHAVIOR

THE ROLE OF INFORMANT PERSONALITY AND CONTEXT

Gert Kroes

Dit onderzoek werd mede mogelijk gemaakt dankzij de financiële steun van de *Stichting tot Dienstverlening aan De Waarden*.



English editing: Lee Ann Weeks (Scientific Editing Service)
Cover design: Jetty Antvelink
Printed by: Ponsen & Looijen, Wageningen

ISBN-10: 90-8100083-1-5
ISBN-13: 978-90-8100083-1-5

THE PERCEPTION OF CHILD PROBLEM BEHAVIOR

THE ROLE OF INFORMANT PERSONALITY AND CONTEXT

een wetenschappelijke proeve
op het gebied van de Sociale Wetenschappen

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de Rector Magnificus prof. dr. C.W.P.M. Blom,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op donderdag 23 februari 2006
des namiddags om 3.30 uur precies

door

Gert Kroes

geboren op 16 december 1952
te Genemuiden

Promotores**:**

Prof. dr. J.W. Veerman
Prof. dr. E.E.J. de Bruyn

Manuscriptcommissie:

Prof. dr. R.C.M.E. Engels (voorzitter)
Prof. dr. M.A.G. van Aken (Universiteit Utrecht)
Prof. dr. J.M. Koot (Vrije Universiteit Amsterdam)

CONTENTS

Chapter 1

Introduction

In clinical practice, information about child problem behavior is often provided by informants other than the children themselves. Particularly in the case of young children, parents are generally considered the most important source of information. When treatment for behavior or learning problems is contemplated, professionals such as teachers and therapists also become important sources of information. In fact, these professionals provide the information needed to evaluate behavior problems, establish diagnoses, determine a treatment strategy, and assess treatment progress. Unfortunately, the opinions of informants with regard to the nature and severity of behavior problems can clearly differ. In actual child and youth care practice, moreover, this is regularly the case. And this situation raises the question of just who provides the most accurate and valid information with regard to child problem behavior, which is the topic of the present dissertation.

In Chapter 2 entitled *Reality and bias in the perception of child problem behavior*, an overview of the relevant research literature is presented. First, the similarities and differences in the judgments of informants are considered. The question of just who provides the most accurate information in which situation is discussed along with the question of whether the information provided by different informants can be fruitfully combined or not. The validity of informant judgments is next considered in connection with the personalities of the parents in particular. A number of examples showing the possible bias of parental perceptions of a child's behavior by the personality characteristics of the parents themselves are presented for this purpose. Finally, the different factors that appear to influence the accuracy of child behavior judgments are subsumed within a more general model of social perception — namely, Funder's (1995) Realistic Accuracy Model (RAM) — for further examination and analysis.

In Chapter 3 entitled *Bias in Parental Reports? Maternal Psychopathology and the Reporting of Problem Behavior in Clinic-Referred Children*, the results of our first study of the possible bias of parental perceptions are presented. In this study, the relations between various types of maternal psychopathology and maternal reports of internalizing and externalizing child behavior problems are examined with a sample of 68 boys admitted to a child welfare institution in the Netherlands (*Stichting de Waarden*) for residential or day treatment. In addition to having the mothers rate the behavior of their 6- to 12-year old boys, the teachers and group-care workers were also asked to assess the boys' behavior, to provide a double criterion measure of the boys´ problem behaviors. All of the informants completed

parallel versions of the widely used Child Behavior Checklist (CBCL; Achenbach, 1991). It was hypothesized that, if such maternal psychopathology as depression or anxiety influence maternal perceptions or judgments of child behavior, then an association between maternal psychopathology and maternal ratings of child behavior would *continue* to exist even after control for teacher and group-care worker ratings of the same children. In other words, when more depressed or more anxious mothers still report greater behavior problems for their children than non-depressed or non-anxious mothers after the judgments provided by teachers and group-care workers have been taken into consideration, evidence for maternal depression or anxiety leading to the over-reporting of child behavior problems (i.e., informant bias) may have been found. However, such an interpretation of our findings is hampered by a serious methodological problem. For the association between maternal psychopathology and maternal over-reporting of problem child behaviors to be taken as evidence for an informant bias, the mothers and other informants must evaluate exactly the same child behaviors (Richters, 1992). And such a methodological requirement can only be fulfilled under carefully controlled circumstances (i.e., in the laboratory where greater control over the relevant behaviors is possible than in the field).

In Chapter 4 entitled *The Impact of the Big Five Personality Traits on Reports of Child Behavior Problems by Different Informants*, the results of our next study — which also meets the requirement of having informants evaluate the same child behaviors — are reported. Videotapes of child behaviors recorded in a clinical play group were used to standardize the behaviors to be observed by mothers, teachers, and group-care workers. Ratings of the videotapes by trained observers who were not previously familiar with the children were also collected to provide an independent criterion measure of child problem behavior. The possible impact of the personality characteristics of the teachers and group-care workers on their ratings of child behavior problems was examined in addition to the impact of maternal personality characteristics on maternal ratings of child behavior problems. Finally, the impact of informant personality in general, as opposed to informant psychopathology in particular, was examined by having all of the informants — with the exception of the independent observers — complete a widely used personality inventory to assess such personality traits as neuroticism, extraversion and openness (NEO-FFI; Hoekstra, Ormel, & De Fruyt, 1996). The same 17-minute videotape was thus observed per child by mothers, teachers, group-care workers, and trained observers for 55 children between the ages of 6 and 12 years admitted to a child welfare institution in the Netherlands (*Stichting de Waarden*) for residential or day treatment.

In Chapter 5 entitled *The Perception of Child Behavior Problems: The Role of Acquaintanceship, Informant Personality, and Context*, the results of our third and final study — which elaborates on the findings and conclusions reached in the previous study — are reported. The impact of both situational and informant characteristics on the perception of child problem behavior are considered in light of Funder's (1999) theory of personality

judgment. No attempts have been made as yet — or to our knowledge — to apply the relevant theoretical concepts and findings from the study of personality judgment to the study of child assessment. It is therefore attempted to do this and to re-evaluate some previously contradictory research findings. More specifically, the use of so-called independent observers who are unfamiliar with the children they observe to attain criterion ratings — in order to meet the methodological requirement put forth by Richters (1992) — is questioned. According to the acquaintanceship hypothesis (Funder, 1999), informants who are familiar with the subjects they observe generally provide more accurate evaluations of their behavior than informants who are not familiar with the subjects they observe (i.e., independent observers). To investigate the role of acquaintanceship and explain some previously contradictory findings, thus, mothers and group-care workers were asked to rate videotaped behavior samples for both a familiar and an unfamiliar child, and these ratings were then compared. Independent observers also rated the same videotapes, and their ratings were compared to the ratings provided by the mothers and group-care workers.

In the study reported in Chapter 5, the role of the context in which the relevant behaviors are sampled is also examined via comparison of the child behavior ratings provided by the mothers and group-care workers for the familiar children in different contexts. More specially, the impact of maternal personality traits on maternal ratings of their children's behavior in everyday life, observed when interacting with their mothers, is examined and compared to the impact of maternal personality on maternal ratings of their children's videotaped behavior samples, recorded in the clinic in the absence of the mothers. We hypothesize that, in everyday life, child behaviors may be actually influenced by such maternal personality traits as neuroticism or extraversion, and therefore an association between maternal traits and maternal child ratings in everyday life may reflect actual child behaviors — rather than personality-related bias in maternal ratings. On the other hand, if an association is found between maternal personality traits and maternal ratings of their children's behaviors recorded in the clinic, in the absence of the mothers, this may be taken as evidence for a bias in maternal reports. The same hypothesis is tested for the group-care workers by examining the relation between group-care worker personality traits and group-care worker ratings of child behaviors during their everyday life experiences with the children, on the one hand, and the relation between group-care worker personality traits and their ratings of videotaped behavior samples of the same children recorded in the absence of the group-care workers, on the other hand. In examining these hypotheses, the issue raised in Chapter 2 of just who provides the most accurate information under which circumstances and the validity of informant judgments in light of their own personality characteristics is again raised.

In Chapter 6, the findings of the three studies of reality and bias in the perception of child behavior problems are summarized. A general discussion of the findings and the implications of the findings for clinical practice and future research is then presented.

In closing, it should be noted that the present dissertation is a compilation of papers. The first three papers have been published, and the fourth paper is currently under review. Such publication inevitably leads to a certain degree of overlap between the introductory and methods sections for the different chapters, but the majority of the data presented in the different chapters is nevertheless original. The focus of each chapter is on a different aspect of the perception of child problem behavior, moreover.

Chapter 2

Reality and bias in the evaluation of child problem behavior[1]

Gert Kroes, Jan W. Veerman, and Eric E. J. De Bruyn

Information about children's behavioral and emotional problems is commonly provided by such different types of informants as parents, teachers, mental health workers, and the children themselves. The agreement between the judgments of different types of informants has been found to be quite low. Different types of informants may be exposed to child behavior under very different circumstances, and behavior problems are often seen to be situation specific. This is one possible explanation for the low agreement between different types of informants. Other possible causes of differences in the judgments of informants may be the characteristics of the informants themselves. That is, information on child behavior may also be informant specific. There is considerable debate about the influence of parental psychopathology on judgments of child behavior, and particularly maternal depression may bias parents' judgments of their children's behavior. It is argued that the evidence for such a bias is weak, at best, due to the lack of clearly validated criteria to evaluate parental judgments. Each informant has his or her own perspective on the behavior of a child, and the best way to proceed is thus to gather as much information of relevance as possible from different sources. Several procedures have been proposed to structure the information attained with regard to the behavior of a child from different sources, but none of the methods has proved superior as yet. The behavior displayed by a child can be treated as a manifestation of one or the other underlying psychological characteristic or personality trait and, in order to improve the assessment of such personality traits, theories of personality judgment and social information processing can be applied. Use of the Realistic Accuracy Model (Funder, 1995), which integrates both situational and informant factors, is recommended to better understand the evaluation of children's problem behavior.

---

[1] This chapter is an adapted version of the Dutch publication entitled *Realiteit en vertekening bij het beoordelen van probleemgedrag bij kinderen* (Kroes, Veerman, & de Bruyn, 2000). English translation: Lee Ann Weeks.

INTRODUCTION

A boy of 10 is brought to an institution for child and adolescent mental health care by his mother. The mother thinks that her son is behaving strangely, is fearful, and hears threatening voices increasingly more frequently. She, herself, suffers from a manic-depressive disorder for which she must regularly be admitted to a psychiatric clinic. She is afraid that her son is displaying the first signs of the same psychiatric disorder. The boy is examined by an intake team and admitted for residential treatment. He initially displays the same symptoms mentioned by his mother. However, the symptoms quickly disappear following admission. During treatment, the boy brightens up, is less anxious, and no longer hears voices. Things are going well at school as well. The boy only falls back into old patterns after weekends he spends at home and even more so after vacations. The difference from how he “normally” behaves is then so marked that it is decided to undertake observation in the home. Although the observers are prepared to see very different behavior on the part of the boy, they are still surprised by what they see. At home, in interaction with his mother, the boy displays exactly the behavior described by his mother at admission. He talks incoherently, says that he hears voices, and clings to his mother. The mother says that her son adapts to the living group and school and therefore does not show his “real” self. The therapists do not completely agree with this interpretation but nevertheless give the mother the benefit of the doubt. In consultation with the mother and a previously consulted child psychiatrist, it is decided to admit the boy to a child psychiatric clinic for observation.

In clinical practice, parents, therapists, and teachers provide important information on the behavior and experiences of children. In fact, these people provide the information needed to evaluate behavior problems, establish diagnoses, determine a treatment strategy, and evaluate treatment progress. Registration with respect to policy objectives and to measure effects also depends on these sources of information to a very large extent. As indicated by the foregoing example, the opinions of informants with regard to the nature and severity of behavior problems can clearly differ. In actual child and adolescent mental health care practice, moreover, this is regularly found to be the case. The question of who provides the most precise and reliable information regarding the behavior of a child is therefore very relevant for clinical practice, policy, and research. And this also explains the large number of studies now available within this domain. The information provided by parents and mothers,

in particular, has received considerable attention. This is because mothers typically have the most intense and long-term relationship with a child and are therefore in a position to provide information that others do not have access to. But what are we to do when the mother herself — as illustrated in the above example — has mental health problems? This is also a situation encountered quite frequently in actual child and adolescent mental health care practice. The question, then, is whether such a situation leads to biased perceptions of child behavior or whether one should speak, as in the preceding example, of accurate perception on the part of the mother but child behavior that clearly depends on the particular situation and appears to be prompted at least in part by the behavior of the mother.

These and related matters will be considered in the present chapter. First, the similarities and differences in the judgments of informants will be considered to the extent that they are relevant for clinical practice. The question of just who, in which situation, provides the most accurate information will be discussed along with the question of whether information from different informants can be fruitfully combined or not. It will be seen that the accuracy of judgments for behavior problems depends upon not only the *situation* in which the behavior presents itself but also the *personality of the informant*. Next, the question of the accuracy of the judgments of informants will be specifically considered in connection with the personalities of parents. A number of examples from research on the possible bias of parental perceptions by the personality characteristics of the parents themselves will be presented. Finally, a proposal to subsume the different factors that can influence the accuracy of the information provided with regard to the behavior of children within a single model will be made. The accuracy of judgments of human behavior is a topic that has received considerable attention in not only the domains of social psychology but also personality psychology. As will be seen, theory with regard to the processes that can influence our judgments of the behavior of others can help us study the validity of the judgments of different informants. In the model we describe, thus, information from a number of different research domains is integrated.

DIFFERENT INFORMANTS

The results of research on the similarities and differences in the judgments of different informants can be globally summarized in the form of three basic conclusions. First, the correspondence between different informants is usually small. Second, every informant has his or her own specific perspective on the behavior or a part of the behavior of a child. Third, different strategies have been developed to combine information from multiple informants but complicated procedures do not appear to work any better than simple procedures for allowing

the judgments of every informant to count. In the following subsections, we will further substantiate these conclusions.

Similarities and differences in the judgments of informants

Information on the behavior of a child can be gathered in a variety of manners. In clinical practice, use is made of standardized tests, interviews, behavioral checklists or questionnaires, and behavioral observations. These instruments lend themselves to the comparison of different informants, and reporting has indeed occurred most with regard to such instruments. The correspondence between informants can occur in more than one manner. One can agree or not agree on the presence of certain clinical characteristics or symptoms. When a particular combination of symptoms is perceived by the different informants, moreover, one can speak of agreement on the clinical syndrome. Research on the similarities and differences in the judgments of informants typically involves having several informants judge the behavior of one and the same group of children. To determine the degree of agreement, the level of the scores may first be examined. And to do this, the average score for each informant is typically used. Average scores are also often taken to provide an indicator of the severity of the problems for a group of children. The degree of agreement between informants can also be determined by comparison of the rank orders of the scores assigned by different informants (i.e., which children are rated as having more problems and which children are rated as having fewer or less severe problems). A large degree of agreement in the rank orders of the scores provided by the informants is usually, then, expressed by a large correlation coefficient, which shows the informants to agree on the relative severity of the observed problems independent of possible differences in the average levels of their judgments. Parents, for example, may generally assign higher scores on behavioral checklists or questionnaires than teachers. When the parents consistently assign higher scores to the same children as the teachers, however, a high correlation is the result — irrelevant of the differences in the average score levels. The reverse is also conceivable, namely that parents and teachers assign a group of children equally high scores in general but completely disagree on the relative severity of the problems for the different children, which results in equal group averages but a low correlation between the individual scores provided by the parents and teachers.

In the following, an overview of research examining the similarities and differences in the judgments of informants with the aid of behavioral checklists and questionnaires will be presented. Level studies will first be considered and then studies in which any rank-order differences in the judgments of informants are examined. Thereafter, a brief discussion of the research examining the similarities and differences in the judgments of informants on the basis of structured interviews will be presented. As will be seen, the interview research

involves primarily level studies. A brief summary of what is known about the similarities and differences in the judgments of informants will then be provided.

Research with the aid of behavioral checklists and questionnaires

LEVEL STUDIES.

In a variety of studies, the average levels of the scores provided by several informants using comparable measurement instruments are reported. A systematic overview of these studies is not as yet available but, within the framework of their own research on the agreement between informants, Meijer and Veerman (1989) have summarized the information from a number of studies. The following points stand out in particular:

- Parents generally report greater problem behavior than teachers (van der Doef & Veerman, 1985; Verhulst & Akkerhuis, 1987).
- Parents and particularly mothers generally report greater problem behavior than group-care workers or other mental health workers for younger children (6-9 years) (Kazdin & Bass, 1988); for older children (14-16 years), mental health workers tend to report more problems than parents (Stavrakaki, Vargo, Roberts & Boodoosingh, 1987).
- The studies are divided with respect to the differences between the parents and the children themselves; sometimes the parents report more problems, sometimes the children report more problems, and sometimes no differences are detected (Hofman & Tates, 1987; Mokros et al., 1987; Stavrakaki et al., 1987; Treiber & Mabe, 1987; Verhulst et al., 1987).

The subsequent research by Meijer and Veerman (1989) in a child psychiatric setting generally confirms the previously detected differences between parents and group-care workers. Parents were found to report more behavior problems than group-care workers but only for children being treated on an outpatient basis. For children receiving residential treatment, the score levels for the parents and group-care workers were equal. Jansen and Oud (1993) also compared the levels of the scores provided on behavioral checklists and questionnaires by parents and group-care workers for a number of adolescent mental health institutes and found parents to report more behavior problems than group-care workers shortly after the admission of a child. Six months following admission, however, the differences in the levels of the scores had largely disappeared. Kroes and van der Doef (1994) observed the same large difference in the level of the scores provided by parents versus group-care workers at the start of treatment. During the course of treatment, the differences decreased not only between the parents and group-care workers but also between the parents and teachers. Parents report less problem behavior as the course of treatment progresses while the scores provided by group-care workers and teachers tend to increase slightly during the course of

treatment to be virtually equal to the level of the scores provided by parents at the end of treatment. Ten Brink (1998) compared the levels of the scores for parents and group-care workers at the beginning and end of treatment within a child psychiatric setting. Large differences were again found to occur in the scores at the beginning of treatment and to disappear by departure. In their research in outpatient treatment centers, Kloosterman and Veerman (1999) similarly found parents to judge the behavior of their children as more deviant at the beginning of treatment than teachers.

In short, level differences exist in the behavioral judgments of parents, mental health workers, and teachers but the size of the differences clearly varies depending on the age of the children, the form of treatment, and the point at which the information is provided (i.e., at the beginning or end of treatment).

RANK-ORDER STUDIES.

In 1987, Achenbach, McConaughly, and Howell published a meta-analysis that still holds as the most complete overview of studies in which empirical information on the agreement between informants in terms of the rank order of their judgments is reported. Achenbach et al. analyzed a total of 119 studies in which different informants judged the problems of children using behavioral checklists or questionnaires. The meta-analysis encompassed a wide range of instruments: questionnaires intended to measure anxiety, depression, rage, assertiveness, temperament, self-control, popularity, or behavior problems in general. The ages of the children varied from 18 months to 18 years. The informants were parents, the children themselves, peers, teachers, mental health workers, and trained observers.

Achenbach et al. (1987) calculated the average correlations across all studies for all possible combinations of informants. Most striking was that the correlations between the informants showed a very consistent pattern: A clear difference in the agreement between the different types of informants. That is, informants with the same type of relation to a child — mothers and fathers, two group-care workers, different teachers — generally agreed to a reasonable extent on the severity of the problems (with correlations varying from .54 to .64). In contrast, the associations between dissimilar types of informants with different relations to the child — such as fathers and teachers or teachers and group-care workers — were generally much lower (with correlations varying from .24 to .42). This pattern repeated itself irrelevant of the instruments used or the more specific combination of informants. Furthermore, the average correlations between the judgments of the children themselves and those of the other informants were lowest (varying from .20 to .27). The sex of the children did not affect the degree of agreement in the judgments provided by the different informants but the age of the children did. The average correlations for 6- to 11-year olds were significantly higher than the average correlations for 12- to 19-year olds. Whether the scores

of mothers versus fathers were correlated with the scores of other informants did not make a difference. Furthermore, the pattern of the correlations for groups of children requiring mental health care was similar to the pattern of correlations for children from the normal population. There was nevertheless greater agreement on externalizing behavior problems (e.g., aggression, acting out) than on internalizing problems (e.g., anxiety, depression, lack of self-esteem). The latter findings did not hold for mothers and fathers, however, who produced exactly the same pattern of correlations for both externalizing and internalizing problems.

The assertion that the agreement between different informants is — at best — small must be refined. Informants with a similar relationship to the child tend to agree to a reasonable extent. Such informants interact with the child under similar circumstances and thus have access to similar sorts of information with regard to the behavior of the child. In contrast, informants with differing relations to the child interact with the child under different circumstances and therefore see different behavior on the part of the child. Different informants can thus have very different perspectives on the behavior problems of children. Achenbach et al. (1987) conclude that the information regarding behavior problems is specific to a particular situation or particular informant. All informants therefore have, according to Achenbach et al., their own value when it comes to the evaluation of behavior problems. As already noted, the observed connections are very consistent and manifest themselves with the use of different types of behavioral checklists and questionnaires. The test-retest reliability of the judgments of the different informants have also been generally found to be reasonable to good. This finding justifies the conclusion, also according to Achenbach et al., that the observed differences cannot be attributed to unreliable measurement methods.

The most prominent findings from Achenbach et al. (1987) have been confirmed in a variety of subsequent studies (Phares, Compas, & Howell, 1989; Hinshaw, Han, Erhardt, & Huber, 1992; Stanger & Lewis, 1993; Sourander & Piha, 1997). Somewhat deviant results are reported by McCombs Thomas, Forehand, Armistead, Wierson, and Fauber (1999) who examined the agreement between different informants for 11- to 15-year olds and found equally high correlations between parents and children for internalizing problems as for externalizing problems. The authors suspect that the explanation lies in the population of children studied and refer to research on the development of the different dimensions of personality (also see Elphick, Slotboom, & Mervielde, 2000). As children get older, their emotional (i.e., internalizing) problems show greater similarities to the emotional problems of adults and are therefore evaluated better by parents as well. The result is greater agreement between the judgments provided by parents and children. For externalizing behavior problems, just the opposite pattern of development appears to occur: Young children display hyperactive and oppositional behavior in a much more uninhibited manner than older children. For this reason, both the visibility of such behavior problems and the agreement between informants decline with the age of the child.

The results of Dutch research generally provide the same picture as the results of Achenbach et al. (1987). Meijer and Veerman (1989) found quite low agreement between parents and group-care workers in correlational terms. Verhulst and Akkerhuis (1989) observed relatively low agreement between parents and teachers although the agreement proved larger for externalizing behavior problems relative to internalizing problems and strikingly high for parents and special education teachers (r=.57) as opposed to parents and regular education teachers (r=.30). Ten Brink (1998) similarly found rather low agreement between parents and group-care workers for externalizing behavior problems at different points in the treatment process (i.e., correlations varying from .24 at admission to .26 at discharge) and very low agreement for internalizing problems (i.e., correlations varying from .09 to .02). Kloosterman and Veerman (1999) found a moderate correlation between parents and teachers for the total group of problems at the start of treatment in outpatient treatment centers (.15). When internalizing versus externalizing problems, boys versus girls, and younger versus older children were distinguished, however, the correlations between the judgments provided by informants proved much higher. For externalizing problems among girls, a correlation of .49 was found; for internalizing problems among older youth, a correlation of .49 was found.

Research with the aid of interviews

The use of structured interviews to classify the behavior problems of children is less common in clinical practice than the use of behavioral checklists or questionnaires (Esmeijer, Veerman, ten Brink, & van Leeuwen, 1998). In numerous clinical and epidemiological studies, however, structured interviews are the most important instruments available to us to systematically attain a classification in terms of the Diagnostic and Statistical Manual of mental disorders (DSM; American Psychiatric Association, 1994). Parents and children are then the most frequently consulted source (Kazdin, 1988). The reliability of structured interviews has been generally found to be moderate to good (Esmeijer, Veerman, & van Leeuwen, 1999). More than the outcomes of behavioral checklists or questionnaires, however, the outcomes of structured interviews vary depending on the informants, the type of interview, or the types of problems being considered. Agreement can pertain to determination of the presence of certain symptoms or classification in terms of syndromes according to the DSM criteria. In both cases, the findings diverge and can vary from moderate agreement between the different informants to a total lack of agreement (Kazdin, 1988; Bidault-Russell et al., 1995). A few general conclusions can nevertheless be drawn and found to largely correspond to the findings for the use of behavioral checklists and questionnaires. Similar types of informants are most consistent in the provision of information irrespective of the type of interview (Fergusson & Horwood, 1987; Valla, Bergeron, Breton, Gaudet, & Berthiaume, 1993). When Verhulst and van der Ende (1991) compared the use of totally different data

collection methods — namely, behavioral checklists and questionnaires versus interviews — by various types of informants, moreover, the agreement between similar types of informants was larger than the agreement between different types of informants irrespective of the type of instrument used. Herjanic and Reich (1982) found the use of overtly perceptible behavior to enhance the chances of agreement between informants while the use of information on state, mood, or feelings typically led to little agreement. Parents and children are again found to frequently judge problems differently depending on the age of the child and the type of problem being considered: What a parent judges as problematic may be perceived as quite normal by a child and vice versa (Kazdin, 1988; Silverman & Eisen, 1992).

Summary of similarities and differences in the judgments of informants

In the following, the agreement between the different types of informants and different instruments discussed above will be considered more systematically. The results of a number of studies have been summarized in Table 1.

Table 1

Pearson correlations between scores for different informants using different instruments.

		Parent		Child		Teacher	Group-care worker
		Checklist	*Interview*	*Checklist*	*Interview*	*Checklist*	*Checklist*
Parent	*Checklist*	.59[1]	.70[2]				
	Interview			.27[2]	.33[2]		
Child	*Checklist*	.25[2]	.27[2]	.74[1]	.64[2]		
Teacher	*Checklist*	.09[3]	.49[2]	.20[2]	.14[2]	.64[1]	
Group-care worker	*Checklist*	.08[4]		.27[2]		.34[2]	.54[1]

1. Achenbach et al. (1987); the average child-child correlation is based on test-retest correlations.
2. Verhulst & van der Ende (1991).
3. Verhulst & Akkerhuis (1989), correlation between Child Behavior Checklist (CBCL) and Teacher Rating Form (TRF) Total Problem scores averaged across sex and age groups of 4-5 years and 6-12 years; Kloosterman & Veerman (1996, 1997, 1999), correlation CBCL and TRF Total Problem scores averaged across sex and age groups of 4-11years and 12-18 years, data at admission.
4. Meijer & Veerman (1989), correlation CBCL Total Problem scores, averaged across outpatient and residential treatment for boys 6 to 11 years; ten Brink (1998); correlation CBCL Total Problem scores averaged across boys and girls 4-14 years, data at admission.

In doing this, we limited ourselves to the results of Dutch research supplemented with information derived from Achenbach et al. (1987) on particularly the agreement between similar types of informants. Within this context, it should be noted that the correlations from

Achenbach et al. are all average correlations and thus calculated on the basis of several studies using different checklists and/or questionnaires. The remainder of the correlations in Table 1 come mostly from a single study with the exception of the correlations between parents and teachers and the correlations between parents and group-care workers. The latter correlations were calculated by us on the basis of the data from a number of Dutch studies and are therefore average correlations as well. The correlations between informants from different subgroups can be seen to differ considerably. A number of the differences have already been mentioned in our discussion of the use of behavioral checklists and questionnaires above. The correlations appear, for example, to sometimes differ depending on the sex of the target population, the age of the target population, the nature of the problem behavior, the type of treatment undertaken, and the point at which measurement occurs. It should be further noted that only total score correlations are presented for the questionnaire-questionnaire agreement comparisons, and that a mean is calculated for data originally presented for multiple subgroups.

Inspection of the results presented in Table 1 shows informants who experience the child in the same situation to agree with each other much more often than informants who experience the child in differing situations irrelevant of the measurement instrument used. The degree of agreement found using behavioral checklists or questionnaires was virtually the same as that found using interviews for the majority of the informants.

In sum, the following conclusions can be drawn with respect to the similarities and differences in the judgments of informants.

1. It is repeatedly shown that different informants provide different information with respect to the problem behavior of children. Such information is thus informant specific.
2. The agreement between informants who experience the child in the same situation is much greater than the agreement between informants who experience the child in different situations. The information is thus situation specific.
3. The aforementioned patterns of agreement and disagreement are observed across different instruments — both behavioral checklists/questionnaires and structured interviews — and also across different types of informant pairs — parents and teachers, parents and the children themselves, teachers and group-care workers, etc. It is therefore very unlikely that the observed differences can be attributed to unreliability of the measurement methods. The information is thus not instrument specific.
4. Parents generally perceive more problems on the part of their children than professionals (i.e., group-care workers, teachers).
5. For children under the age of 12 years, there is usually greater rank-order agreement among informants than for children over the age of 12.
6. There is more often rank-order agreement with respect to externalizing behavior problems than internalizing problems.

In the aforementioned studies, use was made of instruments that are generally known to have moderate to good test-retest reliability. This means that repeated administration of a questionnaire to the same informant leads in most cases to the same judgments. The observed differences between informants cannot, thus, be attributed to unreliability of the measurement instruments. It is still possible, however, that the consistent use of different instruments with different types of informants may lead to so-called method variance (see Campbell & Fiske, 1959; Fiske, 1987). That is, variance caused by the use of a specific instrument with one informant but not with the other may obscure our view of any other differences in the judgments of informants. However, the consistency the consistency with which virtually the same patterns of correlations is found for differing pairs of informants using different instruments makes is probable that the observed differences cannot be attributed to the use of different instruments with different informants alone (Achenbach et al., 1987). Whatever the causes of the detected differences may be, we can conclude that the information provided with regard to the behavior problems of a child depends on the type of informant, the situation in which the behavior is perceived, the age of the child, the nature of the behavior, and the observability of the behavior.

How to handle differences between informants

The small degree of agreement found between the judgments of different types of informants with respect to the behavior problems of children presents not only the clinician but also the researcher with a problem. What are we to do when two informants clearly do not agree? Whose judgments should be treated as decisive? In which situation is the problem behavior of the child most apparent? And who is best in a position to judge whether particular behavior is problematic or not? Different perspectives can be adopted with regard to these questions, and two global trends can be identified. The first trend is simply to consider the contributions of the different informants more or less separately from each other. Discussion then concerns only the question of whether the judgments of all informants should be treated as equal or whether the judgments of certain informants may be more valuable than the judgments of other informants for specific types of behavior. Representatives of the second trend attempt to integrate the information from the different informants in order to attain a shared judgment. Variants of the two trends will be briefly considered further below.

Every informant has a unique contribution to make

Achenbach et al. (1987) concluded in their meta-analysis that information regarding the behavior of children is situation and informant specific. On the basis of this conclusion,

they then argued that the opinions of different informants should be considered individually. Consequently, Achenbach (1985) has proposed a multidimensional model to systematically represent the information from different informants, instruments, and domains of child behavior. In doing this, however, the authors do not make a clear distinction between informant and situation. That is, the situation in which the behavior is observed appears to be considered a characteristic of the informant. The model contains five dimensions or axes for the collection of information on a child: perceptions of the parents, perceptions of the teachers, cognitive examination, physical examination, and clinical examination. For each dimension of information, a number of instruments are available. The model is aimed at the collection of as much information of relevance in the most systematic and reliable manner without striving to achieve a single, unequivocal classification of the child's problem behavior.

A number of authors have refined the view that different sources of information are generally of equal value for the evaluation of behavior problems (Loeber, Green, Lahey, & Stouthamer-Loeber, 1989; Hart, Lahey, Loeber, & Hanson, 1994). Loeber et al. (1989) showed certain informants to provide more useful information with regard to particular types of problems than others. And this is not an unfamiliar phenomenon in clinical practice. When Loeber, Green, and Lahey (1990) examined the diagnostic strategies of a number of clinicians, , for example, they found the information from different sources to indeed be weighted differently. Parents and children are considered better than other informants when it comes to the evaluation of internalizing problems such as anxiety, low self-esteem, or depression. Teachers have the best view of peer-interaction problems. Mothers appear to be more capable of spotting severe behavior problems (i.e., conduct disorders) than other observers of their children. And both mothers and teachers are better informants with regard to oppositional behavior problems than children. Loeber et al. (1989) also therefore plead for the use of different informants without treating them as necessarily equal. Depending on the nature of the problem, that is, different types of informants appear to provide the most useful information.

Combining information from different sources

The purpose of combining the perspectives of different informants is to make the information with regard to a particular child easier to inspect and thereby attain an unequivocal judgment. This may be done for diagnostic or policy reasons, or in epidemiological research. A prerequisite is obviously that no essential information be lost during the process of combining the information. A number of formulas have been tested to combine information from different informants and stated somewhat simply, the methods involve either simple or complicated combination procedures.

One simple procedure for the identification of problem behavior is as follows: if one of the informants ascertains problem behavior, then one can speak of such (i.e., application of the so-called “or” rule). Bird, Gould, and Staghezza (1992) tested this procedure with the aid of the Diagnostic Interview Schedule for Children (DISC; Costello et al., 1987). Diagnoses based on the information provided by parents and children were compared to diagnoses made by clinicians on the basis of all of the information provided by the parents, children, and also teachers. In addition to application of the “or” rule to determine whether parents and children alone provide sufficient information to establish an accurate diagnosis, the authors applied more complicated combination techniques such as multiple regression analyses with the judgments of the clinicians as the criterion, and Loeber’s conditional probability approach (Loeber et al., 1989). The latter two techniques essentially involve identification of those items from the interviews with the parents and children that relate most to the judgments provided by the clinicians. Bird et al. found use of the “or” rule as a combination strategy to go virtually unsurpassed by use of more complicated combination techniques. That is, the combination of information from parents and children produced virtually the same number of accurate diagnoses — with “accurate” meaning correspondence to the clinical judgment in this case — no matter which combination strategy was used. For each of the combination methods, moreover, more problem cases in accordance with the clinical judgments were detected than when the information from only a single source of information was used. In other words, the use of a combination of information sources is more sensitive than the use of a single source for the detection of problem behavior. Fewer problem cases are missed and fewer so-called “false negatives” thus occur when a combination of sources is used as opposed to just a single source.

The employment of information from multiple informants is thus needed, according to Bird et al., to avoid the nondetection of behavior problems. A major disadvantage of using the “or” rule, however, is that more problem instances tend to be detected than is actually the case. This is due to the generally minimal overlap in the information provided by different informants and the fact that all opinions are allowed to count. The risk is thus that not only accurate diagnoses are generated but also inaccurate diagnoses or so-called “false positives”. Using the “or” rule to combine different sources of information, cases of problem behavior that really do not exist may be identified because the method is not sufficiently specific.

One method that overcomes the aforementioned problem of diagnostic false positives involves application of the so-called “and” rule: A problem exists when two or more informants agree that this is the case. A major objection to the use of this rule is again the risk of too few cases of problem behavior being identified.

Combination of the two aforementioned rules in some manner constituted an obvious alternative, and Cohen, Velez, Kohn, Schwab-Stone, and Johnson (1987) did just this. More specifically, a two-step procedure was proposed to identify behavior problems on the basis of standardized interviews. First, the information from multiple informants is combined using

the “or” rule, which is the most sensitive combination method. If a child is judged to meet one of the DSM criteria for a particular disorder by one of the informants, that criterion is considered to be met. If all of the necessary criteria for a particular disorder are met using this rule, the child is provisionally classified as such. Second, the information deemed most relevant for the classification in question is analyzed in greater detail. Prior to the conduct of the interview, a clinician has identified those items that are particularly relevant for a specific classification and those items that are not (see Costello et al., 1987). All of the parent and child items identified as relevant to a particular classification are combined to form a single scale. The scores for a particular child on these items are summed and this total sum score is next compared to the population mean established in previous epidemiological research involving the structured interview. Children who meet the DSM criteria for a particular classification (step one) and with a sum score at least one standard deviation above the population mean (step two) are considered probable problem cases while children with a sum score) two standard deviations above the population mean are considered certain problem cases.

With the integration of simple methods to combine information from multiple sources, we actually end up with a rather complex combination method. By complicated combination methods within the present context, however, we mean the use of such statistical techniques as regression analysis to determine the best indicators for a DSM classification (Kazdin, 1988). Such techniques were previously mentioned in our description of the research by Bird et al. (1992). For multiple regression analysis, use is made of an external criterion considered representative of a particular classification. The external criterion may be a clinical judgment, referral for child or adolescent mental health care, or criminality at a later age. Via the correlation of different variables — including the responses provided by different informants — with the criterion variable, just which variables relate most strongly to the external criterion can be determined and thereby those variables with the greatest predictive utility. The best combination of variables can then be tested under other circumstances or in other situations in order to find an operational criterion for the identification of particular types of problem behavior. This method has been used for the identification of hyperactivity in particular (Power et al., 1988) but also lends itself to the identification of other problem behaviors. The COM procedure developed in the Netherlands (Mesman Schultiz, 1978) is also an example of the use of such a method. This procedure already has a number of different applications and was originally developed as a tool to indicate a need for removal from the home and referral to a specific form of residential help. On the basis of a review of the research literature, a number of variables judged as relevant to the particular diagnostic need are identified. These may be family circumstances and problem behavior on the part of the children. With the application of statistical analyses, those variables that appear to best predict successful removal from the home are next selected from the initial group of variables.

A major problem associated with the above method is determination of a valid criterion. Who determines the initial diagnosis to be used for comparison to the judgments of others? Who determines what constitutes successful removal from the home? On what grounds? Using which informants? Kazdin (1988) has argued that no clear answer exists for these questions but also sees no reason to be concerned about this as good criteria and good predictors are often sought simultaneously during the development of reliable measurement instruments.

Comparison of combination methods

Drawing upon Bird et al. (1992), Offord et al. (1996) also compared different methods for the combination of information from multiple sources. Using the Revised version of the Diagnostic Interview for Children and Adolescents (DICA-R; Reich & Welner, 1988), information was collected from parents and teachers in order to reach a DSM-III-R classification for a number of psychiatric disorders. The following combination methods were then applied successively: independent classification by every informant; the "or" rule; and then the "and" rule. The classification results were next compared to a number of external criteria including the age and sex of the children, single-parent family, family income, depression on the part of one of the parents, and family problems. The internal consistency of the different methods was first calculated and found to produce comparable scores (i.e., alpha coefficients of .60 to .91). This means that isolated use of the different combination methods led to more or less equally reliable results at least in terms of internal consistency. The correlations with the external criteria were next calculated and found to produce a very different picture. All of the informants and every combination method revealed completely different and unique patterns of associations with the external criteria. The classification of a behavior disorder according to the judgments of the parents, for example, correlated strongly with depression on the part of one of the parents and poor family functioning. When a behavior disorder was identified on the basis of teacher information, no associations with depression on the part of one of the parents or family functioning were detected. This finding led Offord et al. to conclude that no combination of evaluations should be undertaken with respect to psychiatric disorders. Every informant provides specific information that cannot be supplied by others and may be lost to a significant extent with the application of combination methods. As long as we cannot determine which information is most relevant for the identification of a particular disorder with certainty and which information is not, it is best that every informant be respected for what he or she can contribute and the information from different informants thus be given careful and complete consideration.

With the preceding insights, we return to the perspective of Achenbach (1985) and Achenbach et al. (1987), namely that information with regard to children appears to be informant specific and that we should therefore assemble information from multiple sources

preferably in the most systematic manner possible. Offord et al. (1996) share this opinion and argue that psychiatric disorders can best be approached as informant-specific phenomena. For the moment, such an approach offers the only guarantee that relevant information will not be lost and that the perspective of each and every informant will be given due respect. A definite disadvantage of such an approach is the risk of generating a large amount of disordered information. As Offord et al. (1996, p.1084) observe:

Clinicians usually bring data together from different informants in a nonsystematic manner to arrive at a diagnosis. Given our current knowledge and expertise in classification, this is the most practical way in which to proceed in a clinical setting to generate a diagnosis. However, by factoring in the relative contribution of informant-specific information, the clinician may be aided in teasing out questions of etiology, natural history, and response to treatment among his or her patients.

THE ROLE OF INFORMANT CHARACTERISTICS

One of the major conclusions in the preceding was that judgments of child behavior depend to a significant extent upon the situation in which the behavior is perceived and the type of informant. This is often referred to as situation or informant specificity. The terms are frequently used interchangeably and often not very clearly. “Situation specific” typically means that the behavior of the child may occur in one situation but not another. “Informant specific” typically means that one informant but not the other can see certain behavior. The term “informant specific” is also often used to mean that the informant is in the situation in which the behavior presents itself. However, “informant specific” in the latter sense does not really differ from “situation specific.” A teacher (i.e., a specific informant) generally experiences a child at school (i.e., within a specific situation). Is problem behavior displayed in the classroom then: informant specific, situation specific, or both?

In order to better distinguish the influence of situation and informant on the accuracy of judgments of behavior problems, we formulated a new definition of situation and informant. Under situation, we understand the circumstances in which the behavior of a child presents itself. This can be the place in which the behavior occurs, but also the influence of the environment on a particular behavior or the interaction with a specific person that provokes a particular behavior. If a child listens to one teacher but not the other, “the individual in front of the class” is a characteristic of the situation and listening is — in this case — situation specific. Under informant, we understand the person who perceives and evaluates the behavior of the child. Informant specific means: dependent on those personal characteristics and circumstances of the informant that may influence their judgments of child behavior. Informant characteristics in the case of teachers may be, for example: whether the teacher is in a position to notice the behavior (e.g., in cases of surreptitious or sneaky

behavior), whether the teacher is tired, whether the teacher is distracted by others in the class, whether the teacher has experience with children with behavior problems, the expectations that the teacher has with respect to a particular child, and the reference framework utilized by the teacher.

In this section, we will consider the personal characteristics of the informant in greater detail. A number of factors that can influence the judgments of informants will be discussed. In doing this, however, we will limit ourselves to the parents as important informants with regard to the behavior of their children. Kazdin (1988) has identified a number of characteristics that can influence parental judgments: mental health problems, marital problems, expectations with regard to the behavior of their child, self-esteem, and social support from the family and the surrounding social environment. A brief overview of research on two of these characteristics — namely, mental health problems on the part of mothers and the influence of aggressive child behavior on parental perceptions of child behavior — will be presented in the following.

The influence of maternal mental health problems on judgments of their children

For some years now, there has been a lively discussion of the influence of parental mental health problems on judgments of child behavior problems (for a review, see Richters, 1992; for more recent publications, see Briggs-Gowan, Carter, & Schwab-Stone, 1996; Sawyer, Streinder, & Baghust, 1998). A great deal of research has been conducted on the influence of maternal depression in particular. Depressed mothers report greater problem behavior on the part of their children than nondepressed mothers. The discussion is concentrated on the question of whether the relatively high degree of problem behavior actually occurs in reality or is simply a consequence of inaccurate perception on the part of depressed mothers. Depressed mothers often see problems that other informants in other situations do not see. Richters (1992) outlines a number of possible explanations for this. One explanation is that children display behavior problems in the presence of their mothers but not in the presence of others, as illustrated in the introductory example. Another explanation is that depressed mothers actually have greater expertise than other informants with respect to the perception of problem behavior and particularly emotional problems on the parts of their children. Alternatively, it is possible that depressed mothers can handle less frustration than nondepressed mothers and thereby perceive child behavior as more quickly problematic than other mothers. Finally, it is possible that depressed mothers have a tendency to project their problems and negative mood onto their own children. A “frustration model” along the lines outlined above would lead to an overreport of externalizing behavior problems for children

(e.g., “he is so busy,” “he is so rebellious”) while a “projection model” would lead to an overreport of internalizing problems (e.g., “he is so anxious, sad, down”).

Up until now, a clear answer to the question of whether one can speak of inaccurate perceptions of child behavior by depressed mothers or a truly elevated degree of behavior problems in families with a depressed mother has not been available. Different studies claim to have demonstrated the reality of an elevated level of child behavior problems as reported by depressed mothers (Breslau et al., 1988; Richters & Pellegrini, 1989). Chilcoat and Breslau (1997), for example, found not only the elevated incidence of behavior problems in reality to be confirmed by the teachers of certain children but also the overestimation of problems by their depressed mothers. The children of depressed mothers thus had more behavior problems in reality than the children of nondepressed mothers but not as many as their mothers reported. Conrad and Hammen (1989) also found a combined effect: The children of depressed mothers showed relatively greater behavior problems but the interaction with a problematic child was also found to lead to a greater reporting of child behavior problems by depressed mothers than was actually the case. Similarly, Jensen et al. (1988), Fergusson et al. (1993), and Peters (1998) observed a connection between a heightened reporting of child behavior problems and multiple forms of psychopathology on the part of mothers including maternal depression. The aforementioned associations and results considered together clearly imply two things: 1) in families where the mother has mental health problems, the children also tend to show relatively more emotional and/or behavior problems; 2) mothers with mental health problems in some instances tend to perceive more problems on the parts of their children than is actually the case.

Richters (1992) has raised a number of methodological doubts about the design of research on the relation between depression and the presence or absence of child behavior problems. According to Richters, the aforementioned claim that depression has been shown to influence perception is also therefore open to debate. In order to show that depressed mothers perceive more behavior problems on the parts of their children than is actually the case, the judgments of the mothers should be compared to the independent and validated judgments of other informants evaluating comparable child behavior in comparable situations. When depressed mothers are indeed found to perceive more problems than other informants evaluating the same behavior, evidence that the detected overreport of behavior problems by the depressed mothers is caused by their depression and not one of the many other family factors that can possibly influence such perceptions must also then be supplied. Richters analyzed 22 studies of the influence of depression on the part of mothers on the reporting of child behavior problems and found none of the studies to meet the aforementioned requirement. The second requirement constituted the most important obstacle: namely, comparison of the judgments of the mothers with the carefully validated criterion judgments of other, independent informants. As noted in the preceding section, comparison of the judgments of different informants — which were mostly mothers and teachers in the studies

analyzed by Richters — consistently shows very little agreement in terms of both the level of scores and the rank order of scores. This discrepancy is usually attributed to the difference in the situation. Teachers have access to other information than parents (i.e., mothers), and children behave differently in different situations. These differences can then obscure our view of the influence of such informant characteristics as maternal depression on evaluations of child behavior. An elevated reporting of child behavior problems by depressed mothers certainly cannot, thus, be taken as direct proof that such mothers tend to overestimate the incidence or severity of behavior problems on the parts of their children. In order to study the influence of informant characteristics such as maternal depression, the situation must be held as constant as possible or — in other words — the first requirement outlined by Richters must be met. That is, mothers and other informants must clearly evaluate comparable child behavior under comparable circumstances in order to control for situation characteristics. Unfortunately, researchers have rarely managed to do this up until now (Briggs-Gowan, Carter, & Schwab-Stone, 1996).

The influence of aggressive child behavior on parental judgments

Research on the influence of mental health problems on the part of the mother on her evaluation of the behavior of her child seeks the cause of possible bias in the character of the informant. In research on the influence of aggressive child behavior, in contrast, the cause of such bias is sought in the character of the child. It is assumed that the mothers of aggressive and otherwise troublesome children interpret the behavior of their children much more negatively than the mothers of normal children (Dix & Lochman, 1990; Strassberg, 1995, 1997). That such mothers interpret the troublesome behavior of their children more negatively is not illogical as they can certainly speak from experience. But it appears that the mothers of problem children also interpret apparently neutral or even cooperative behavior on the part of their children in a more negative manner than the mothers of normal children. A consequence of perceived negative behavior is often punishment or aggressive behavior on the part of parents. In such a manner, the interaction comes full circle: punishment can lead to further child aggression and the children thereby confirm negative parental perceptions (Patterson, 1982, 1986). The biased perceptions of mothers in the form of viewing neutral behavior as hostile is considered an important link in the chain of negative interactions that can persist between mother and child. It is assumed — in keeping with theory regarding the processing of social information from Dodge (1993) — that social experiences constitute a type of "database" that permeates all future information (e.g., "once a thief, always a thief"). The slightest systematic deviation in the processing of social information can thus have major consequences for the course of social interactions (Orobio de Castro & Bosch, 1997).

Research on the accuracy of maternal perceptions of aggressive child behavior has produced conflicting results (Kendziora & O'Leary, 1998). The problems that such researchers encounter are very similar to the problems that we saw for attempts to study the influence of maternal depression. Richters (1992) is again referred to indicate the shortcomings of research on the perceptions of aggressive child behavior. Demonstrating that the mothers of aggressive children more frequently or quickly call the behavior of their children hostile than the mothers of normal children is the first step to be taken. The question, however, is whether the elevated reporting of aggressive behavior is a consequence of many years' experience with a difficult child (with the child as the cause, thus) or proof that the mother of an aggressive child tends to perceive her child's behavior in a biased manner (with the mother as the cause, thus). When the latter is the case, a vicious circle emerges in which the negative reactions of the child confirm the negative expectations of the mother in the form of a "self-fulfilling prophecy" (Darley & Fazio, 1980). Furthermore, no specific or clear cause can be pointed to and the interactions of mother and child simply reinforce each other. An additional problem associated with this type of research is that it is often conducted on the basis of video recordings of situations involving actors or pictures depicting certain behaviors. Parents are then asked to imagine the child in the video or picture as their own child. The problem, then, is that we no longer know which information the parents evaluate: The behavior of the actors in the video or the behavior of their own child in the past.

A typical example of such research is that of Strassberg (1995). Strassberg had the mothers of children with behavior problems and the mothers of normal children both view video recordings of a play situation involving a mother and a child. Both the mother and the child were actors and behaved as dictated by a script. Each mother was asked to imagine that the video was of her son with her, and identification was promoted by regularly using the name of the viewing mother's son during the play of the video. The child in the video displayed a number of behaviors that were categorized as "clear" (e.g., clear obedience or overt opposition) or "ambiguous" (e.g., negotiation with regard to cleaning up, or whining) by the researchers. The mothers who watched the video were then asked to evaluate the different behaviors; about their emotional reactions to the behavior of the child ("Would you be angry at your son?", "Would the behavior of your son make you nervous or worried?"), about their expectations with regard to the boy's behavior ("Will he quit protesting later?"), and about their evaluation of the child's intentions ("Do you think that your son would do this on purpose?"). The reasoning behind the design of Strassberg's study was as follows. Ambiguous situations lend themselves to the projection of one's experiences more than clear situations (Dodge, 1986). On the basis of their experiences with a difficult child, the mothers of aggressive children could be expected to express more negative emotions and attribute more negative intentions to the child in ambiguous situations than the mothers of normal children. In clear situations, in contrast, the mothers of aggressive children could be expected

to rely upon the depicted behavior more and their reactions therefore not differ as much from the reactions of the mothers of normal children.

Strassberg (1995) found confirmation of the aforementioned presuppositions in his study. The mothers of aggressive children generally reported more negative emotions and attributed more hostile intentions to the videotaped child in ambiguous situations than the mothers of normal children. In addition to this, the mothers of aggressive children worried more about the behavior of the child and had more negative expectations with regard to the outcome of the interaction in the video in all of the situations and thus in those situations in which the behavior to be evaluated was quite clearly positive or neutral as well. This finding again elicits the question of which mechanism is precisely in operation during the perception and evaluation of behavior by the mothers of aggressive children. Strassberg again refers to Dodge (1986) who suggests that social-information processing occurs in five steps. The initial two steps pertain to the perception and interpretation of social information (i.e., encoding and representation). The last three steps pertain to the preparation and provision of a response (i.e., response search, response decision, and enactment). The first two steps in the present case are most relevant. Encoding concerns the sensory perception of social information and the storage of such in memory. While perception necessarily involves a certain degree of selection for information that is relevant or irrelevant, the actual interpretation and evaluation of incoming information only takes place in step two. Via the weighting of information, reasoning, and comparison with information that is known to us from previously experienced situations, we determine what exactly the incoming information means for us. Strassberg applied these notions in his own research and concluded that there are two possible explanations for the results that he found. On the one hand, the mothers of aggressive children may accurately perceive the information in the video (i.e., encode correctly) but interpret the information in a negative manner (i.e., on the basis of their experiences with an aggressive child), which explains the negative reactions in ambiguous situations. On the other hand, the mothers of aggressive children may have an inaccurate and biased perception of the information in the video to start with, which means that the encoding of the behavior observed in the video in clear situations is already different than that by mothers of normal children. Strassberg closes with the conclusion that research in which both encoding and interpretation are measured is necessary to better understand the observed differences in the processing of social information by the mothers of aggressive versus normal children.

Both the encoding and interpretation of information regarding social behavior are processes that occur within the informant. The discrimination of the two processes may thus enlighten research on the role of informant characteristics in the evaluation of child behavior. With regard to the example from the research by Strassberg (1995), we may add that the differing situation characteristics created extra confusion. That is, the instruction “imagine finding yourself in the situation from the video with your own child” makes the “clear situations” unclear. Just which behavior is encoded and subsequently interpreted by the

mothers of aggressive children? The behavior of their own aggressive child in the past or the neutral behavior of the child in the video? When the mothers of aggressive children perceive hostile behavior in clearly neutral situations, is their encoding incorrect or are they literally following the instructions they were given? Strassberg also recognized this problem and therefore suggested that parents should be asked to evaluate the videotaped behavior of both their own child and other normal children in future research. In such a manner, whether the parents of aggressive children judge the behavior of children more negatively in general and thus irrespective of actual observed behavior (or not) should become clear. Kendziora and O'Leary (1998) did just this when they compared the judgments provided by the parents of problem children for their own child versus other children. The same was also then done with the parents of normal children. The parents of problem children appeared — just as the parents of normal children — to judge the behavior of their own child more positively than the behavior of other children. Mothers even labeled the behavior of children as negative or off-limits less often in general than trained observers within the same study. A clear bias was thus detected on the part of mothers but in a totally different direction than expected.

Just as for research regarding the influence of depression on the evaluation of the problem behavior of children, research regarding the influence of aggressive child behavior on the judgments of parents shows the importance of carefully distinguishing characteristics of the situation from characteristics of the informant. In the following closing section, we will incorporate both situation and informant characteristics into a single model in an attempt to gain greater insight into the influence of both types of characteristics on judgments of child behavior.

AN INTEGRATED MODEL

The accuracy of judgments of human behavior is a topic that has received considerable attention within the domains of social and personality psychology. Theory regarding the processes that determine the accuracy of our judgments regarding personality characteristics can be of excellent use for research on the evaluation of child behavior problems by different informants. More specifically, we propose an approach based on the Realistic Accuracy Model of Funder (1995), which includes "attempts to identify the psychological properties of people, such as personality traits" (p. 652). People are continually busy judging others — and themselves. People do this in order to estimate how they can best behave, what they can expect from others, what caused an event to unfold as it unfolded, and so forth. The term "personality traits" refers to not only complex personality dimensions such as extraversion, depression, openness, and such but also "everyday" properties such as friendliness, sloppiness, easily angered, energetic, pushy, sad, disobedient, nice, and so forth

(Funder, 1991). These psychological characteristics are not directly perceptible. Anger can be deduced from physical cues usually in combination with an evaluation of the situation in which the cues are seen. Sloppiness is a characteristic that is not itself perceptible but certainly deducible from all kinds of behaviors such as letting things lie about, forgetting appointments, not combing one's hair, and so forth. Whether or not we are correct in concluding that someone is sloppy must, in the end, be seen. The starting assumption underlying Funder's Realistic Accuracy Model is that personality characteristics or traits are real existing properties of people. In the identification of such on the basis of outward behavior, we can be right or wrong although this is very difficult to prove because the characteristic itself is real but imperceptible. Every term that we use to refer to a human characteristic, moreover, "intends to describe two real things: a pattern of behavior and an inferred attribute of the person who performed it" (Funder, 1991, 1995, p. 653).

In the evaluation of problematic child behavior, the same two things are of concern: outwardly perceptible behavior and the presupposed psychological characteristics or traits expressed by the behavior. Sloppy, busy, aggressive, nice, anxious, sad, cheerful, concentrated, easily distracted, chaotic, insecure, or obedient are all terms used to refer to psychological traits of the child and those behaviors that are an expression of such. In judgments of child behavior, as discussed throughout this chapter, evaluation of the psychological traits of the child is nevertheless of principal concern. Judgments of outwardly perceptible behaviors simply constitute a tool for this purpose — no matter how important such judgments may be in their own right. A few examples may illustrate this. We observe a child crying; for our reaction to this behavior, however, it is essential that we know whether the crying is an expression of grief, happiness, or anger. We see a child running in front of a dog; is this a sign of fear or part of a game? Heated discussions can arise with respect to the question of whether a child hits someone on purpose or accidentally bumps into someone. Similarly, not listening to a parent or teacher may be an expression of opposition or simply due to distraction. The evaluation of behavior does not depend upon solely the perception of physical expressions — although this is a very real part of the evaluation of behavior — but on the identification of psychological traits in order to be able to predict future behavior and act upon this information. The preceding examples also show the identification of personality traits to typically not be based upon a single perception but on a number of behaviors that provide cues to the presence of a particular trait. It nevertheless remains the case that we identify the traits of a child on the basis of concrete behavior in principle. Whether our assessment of personality characteristics is correct or not cannot be determined with certainty. What we can do is strive to approach reality as closely as possible. This can be done by assembling as much information about the child and the circumstances under which various traits manifest themselves as possible — and also as much information as possible about the informant, his or her power of discernment, and the circumstances under which his or her

judgments are made. Funder (1995) describes the assessment of personality characteristics in the following manner.

Description of the model

The process of assessing the personality characteristics of an individual can be schematically outlined as in Figure 1.

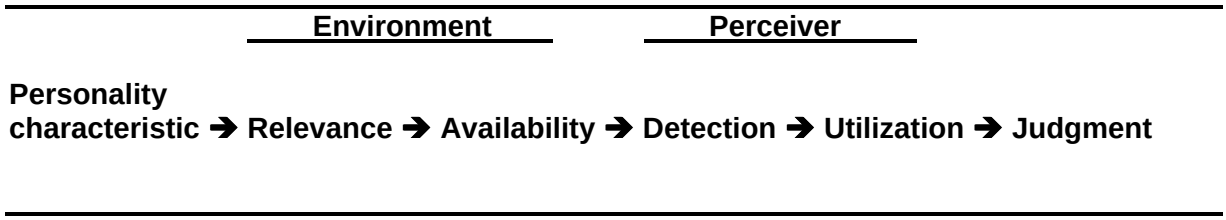


Figure 1. The assessment of personality characteristics according to the Realistic Accuracy Model (Funder, 1995).

The assessment of personality characteristics occurs in four steps. The objective of the process is to identify an — as yet unknown — trait of a person as accurately as possible. First, a personality trait gives rise to certain behaviors under certain circumstances. Such behaviors are referred to by Funder (1995) as relevant to the trait. Second, the behaviors must be perceptible or available to the evaluator. A propensity to steal, for example, is a trait of a thief; the behavior that this gives rise to, however, does not occur very frequently. Third, the relevant and available behavior must be discovered or detected by the evaluator. If a thief steals, for example, this must be noticed. This step is often difficult as the evaluator may be inattentive, simply not present, or the behavior itself may be fleeting and thus difficult to perceive. Fourth, the relevant, available, and detected behavior must be interpreted or utilized properly. It stands to reason that this step can also easily lead to err: The person who sees a man walking with a racing bike over his shoulder can think that the man is training for a cross-country race, that the man has lost his key, or that the man is stealing a bike.

It should be noted that the process of assessing a single expression of a particular personality characteristic or trait is outlined in the diagram in Figure 1. The identification of personality characteristics usually requires multiple observations, which are then compared and combined. To avoid unnecessary complication of the diagram, however, the description of a few evaluations is deemed sufficient. And different methods for the combination of information can be used, as discussed earlier in this chapter. None of these are particularly preferable, and Funder (1995) does not provide guidelines with regard to the selection of such.

The sequence of successive steps outlined above implies that the assessment of personality characteristics is a cumulative process. If one of the links is weak, then the accuracy of the assessment declines accordingly. And that which holds for a single judge also — of course — holds to an even stronger extent when the assessment is made by a combination of judges.

Application of the model

To illustrate the application of the model, we will examine the different steps in connection with a number of examples of the evaluation of problematic child behavior. The first two steps concern the characteristics of the situation in which the behavior occurs. Step one or identification of the relevant aspects of problematic behavior plays, for example, an important role in the research of Strassberg (1995). Some behaviors may be less unequivocal representatives of a personality characteristic than other behaviors simply because the behaviors are relevant to more than one trait. “Negotiation”, for example, may be a relevant expression of “active problem solving” and thereby indicate social competence, but it can also be an expression of “passive resistance” and thereby relevant for the identification of an oppositional personality trait on the part of the child. Similarly, “blunt refusal” can be perceived as assertive behavior (i.e., social competence), or defiant behavior (i.e., oppositional attitude). The clearer the relevance of a particular behavior, the greater the chances of accurate evaluation and judgment. Conversely, the less clear the relevance of a particular behavior is, the larger the chances of differences in the interpretations provided by the evaluators.

With regard to the second step, clear perceptibility of relevant behaviors enhance the accuracy of evaluation considerably (Funder & Colvin, 1991; Levesque & Kenny, 1993). This explains why externalizing behavior leads to greater agreement between informants than internalizing behavior. A depressed mood is less easy to deduce from concrete behaviors than rage or aggression. And in the example presented at the beginning of this chapter, tracking down the relevant behaviors for the evaluation of a child who suffered — according to his mother — from a psychiatric disorder constituted an important part of the problem for the diagnosticians. Conversely, behavior that consistently occurs across different situations and varying circumstances is much easier to evaluate. Mental health workers usually see less problem behavior than the parents at the beginning of treatment, which presumably relates to the perceptibility of the behavior to a large extent. Given that children tend to adapt to a new environment, it is possible that behavior problems scarcely manifest themselves to start with and only rear their ugly head at a later point. Little agreement between parents and teachers can thus undoubtedly, at least in part, be traced back to the perceptibility of the behavior in different situations.

Steps three and four concern the perceiver or informant. Step three includes the detection of relevant and perceptible behavior on the part of the child. This is also where step one from the model of Dodge (1986) actually starts — namely, the encoding of social information. Informants may distinguish themselves with regard to the extent to which they have an eye for different behaviors on the part of a child. Similarly, some people are more sensitive to emotional signals than others, with women generally having a slight edge over men (Hall, 1990). This may also explain why mothers often differentiate better than fathers during the report of child emotional problems (Phares, 1997). Personal traits can obscure our view of certain child behaviors, moreover. And depression on the part of mothers can in some cases lead to an oversensitivity to emotional problems (Richters, 1992). Similarly, long-term experience with a difficult to handle child can make a parent oversensitive to signs of aggressive behavior (Strassberg, 1995).

Finally, in step four, the perceived behavior is interpreted. This step corresponds to step two in the model of Dodge (1986). The accuracy of an interpretation is generally enhanced by such informant characteristics as experience, openness, objectivity, intelligence, expertise, and motivation (Funder, 1995). Such personal traits as depression, hostility, rigidity, or defensiveness can negatively influence judgments of child behavior. The influence of depression and hostility were considered in detail in the previous sections. The studies mentioned there, however, have still not succeeded in showing the interpretations of child behavior by mothers with mental health problems to be less accurate than those by other mothers. This is because the accuracy of judgments regarding child behavior problems are the result of a cumulative process with the interpretation of the behavior constituting only the final step.

CONCLUSIONS

In the evaluation of child behavior problems, one is dependent on the judgments of different informants who certainly do not always agree with each other. Different informants see a child in different situations and frequently see different behavior, thus. Behavior that consistently occurs in multiple situations leads to a unanimous judgment more quickly than behavior that is situation bound. Informants who experience the child in comparable situations agree more often with each other than informants who experience the child in divergent situations. Different informants may also have different interests in the evaluation of a child, different reference frameworks and norms, and also different levels of expertise with regard to the evaluation of personality characteristics. A complicating factor is further that children with behavior problems are often judged by parents who, themselves, have psychological problems. As long as the value of the judgments of separate informants cannot be determined with certainty, combining the judgments of different informants to form a

single final judgment provides little guarantee of a better quality evaluation. The loss of information during the combination of information appears to be unavoidable. In short, both situational factors and characteristics of the informant complicate unequivocal and accurate evaluation of behavior problems on the part of children.

The aforementioned state of affairs actually does not differ much from the state of affairs that we encounter on a daily basis when we evaluate the behavior of other people. People are continually judging each other, but the accuracy of their judgments is certainly open to doubt on many occasions and they certainly do not always agree with each other. Nonetheless, these judgments determine how we behave in interaction with others and just how we feel to a very significant extent. The evaluation of others (and ourselves) is therefore a topic of study for many researchers in both the domains of social and personality psychology. Considerable research has also been done on the evaluation of personality characteristics by different informants. The theories developed on the basis of this research may certainly help us tackle the problems associated with the evaluation of child behavior problems. Application of such theories allows us to place the discussion of reality or bias in the evaluation of child behavior within a broader framework and may thereby help explain the similarities and differences observed between informants. The Realistic Accuracy Model of Funder (1995) provides such a framework because the influences of the situation and the informant are clearly distinguished but also related to each other. The model makes what we already knew even more abundantly clear, namely that accurate assessment of underlying child personality traits is a difficult task and that poor agreement between informants is rather the rule than the exception. At the same time, the model shows what we can do to allow judgments regarding a child to mirror actual traits as closely as possible. In sum, use of the present model may not provide the most simple method for evaluation of problematic child behavior but hopefully the most realistic method.

Chapter 3

Bias in Parental Reports? Maternal Psychopathology and the Reporting of Problem Behavior in Clinic-Referred Children

Gert Kroes, Jan W. Veerman, and Eric E. J. De Bruyn

In this study we examined the relationship between various types of maternal psychopathology and maternal reports of internalizing and externalizing child behavior problems in a clinical sample of 68 boys ages 6-12 years. The reports of both teachers and group care workers were used as criterion ratings. Multiple regression analysis indicated substantial partial correlations between various types of maternal psychopathology and the reporting of internalizing child behavior problems, the latter, after controlling for variance shared with independent raters. Only small to insignificant partial correlations were found in the case of externalizing behavior problems. Results suggest distortion to be associated primarily with internalizing child behaviors. This finding is consistent with the social attribution theory that predicts greater distortion when observing more ambiguous stimuli.

INTRODUCTION

There is still considerable debate about the influence of maternal psychopathology on the reporting of children's behavior problems (Briggs-Gowan, Carter, & Schwab-Stone,1996; Chilcoat & Breslau, 1997; Mick, Santangelo, Wypij, & Biederman, 2000; Najman et al., 2000; Sawyer, Streiner, & Baghurst, 1998; Youngstrom, Izard, & Ackerman, 1999). The question is how to interpret higher levels of problematic child behavior reported by mothers who themselves display psychopathological symptoms: Do the reports of these mothers reflect actual existing behavior problems in children, or does maternal psychopathology influence the perception of their child's behavior problems? Richters (1992) described these two different interpretations as the *accuracy model* and the *distortion model*.

The accuracy model is supported by numerous high-risk studies demonstrating the link between maternal psychopathology and emotional or behavioral disturbances in children (e.g. Beidel & Turner, 1997; Downey & Coyne, 1990; Goodman & Brumley, 1990; Hammen, Burge, & Stansbury, 1990; Kashani, et al., 1990; Rutter, 1990). Genetic and environmental factors both seem to be involved, but the exact nature of the underlying mechanisms is far from clear. For example, the child's problems might result directly from interacting with a depressed or anxious mother, or from factors caused by maternal psychopathology such as family distress, marital conflict, or child neglect. Whatever the exact mechanisms may be, children of mothers with psychopathological symptoms run a greater risk of developing problematic behaviors. These studies seem to justify the conclusion that mothers with various kinds of psychopathology accurately report higher levels of problematic behavior in their offspring.

On the other hand, many studies investigating the distortion model claim that maternal psychopathology leads to a general over-reporting of symptoms in their children (Richters, 1992). For instance, Jensen, Traylor, Xenakis and Davis (1988) examined the influence of a range of psychiatric symptoms in both mothers and fathers on the reporting of child behavior problems. These authors found a variety of positive associations between parental psychiatric symptoms and their reports of children's behavioral problems. Overall results indicated that parent-parent, parent-teacher, and parent-child agreements about children's behavioral and emotional problems were significantly affected by the parents' own gender, sex of the child, and type of parental symptoms. Mothers' symptoms seemed to more affect their reports of sons than daughters, and had a greater influence on internalizing than on externalizing child behavior problems. Moretti, Fine, Haley, and Marriage (1985) and Kolko and Kazdin (1993) observed parent-child differences in reporting internalizing child symptoms related with

parental psychopathology, whereas Breslau et al. (1988) reported mother-child discrepancies on externalizing behavior problems associated with maternal depression. Only a few studies found no evidence in favor of the distortion hypothesis (e.g. Richters & Pellegrini, 1989; Conrad & Hammen, 1989), or concluded that the size of any bias in mothers' or fathers' reports of childhood behavior problems was likely to be very small and of little clinical significance (Sawyer, Streiner, & Baghurst, 1998).

However, most recent studies reported differentiated results, suggesting that both models might hold true: Parental psychopathology may lead to a real increase in behavior problems in children and to an over-reporting of these problems by parents experiencing psychopathological symptoms (Boyle & Pickles, 1997; Briggs-Gowan et al., 1996; Johnston & Short, 1993; Chilcoat & Breslau, 1997; Fergusson et al., 1993; Mick et al., 2000; Najman et al., 2000). For example, Chilcoat and Breslau (1997), investigated accuracy and/or distortion effects of mother's history of major depression, anxiety disorder or substance abuse/dependence on their reporting of children's internalizing and externalizing behavior problems in a community sample. Teachers were used as criterion informants. Chilcoat and Breslau found that children of mothers with a history of major depression have more internalizing problems according to both mothers and teachers. When comorbidity was taken into account, mothers with symptoms of anxiety, a combination of depression and anxiety, or substance abuse appeared to overstate their children's internalizing and externalizing behavior problems, but no evidence of over-reporting was found when mothers were depressed (but not anxious). Najman et al. (2000) used mother and child reports to evaluate the association between maternal mental health and behavior problems in 14-year old children. They concluded that anxious and/or depressed mothers tend to report more cases of child behavior problems — both internalizing and externalizing — than do their mentally healthy counterparts or children themselves. Mick et al. (2000) investigated the effects of maternal depression on mothers' reports of comorbid major depression in adolescents with and without attention-deficit/hyperactivity disorder (ADHD), also using adolescent self-reports as criterion measure. They found evidence of over-reporting by mothers only in the non-ADHD control group of children (not referred for treatment), suggesting that the potential distortion caused by maternal depression may be stronger in community than in clinical settings.

The use of teachers as criterion informants in the study by Chilcoat and Breslau (1997) was criticized by Biederman, Mick, and Faraone (1998) and Ingersoll and Eist (1998). These authors argue that an alternative explanation to the findings of Chilcoat and Breslau would be equally applicable, i.e. that children of depressed and anxious mothers display more behavior problems at home than at school. Their discussion illustrates the main methodological problem in studies examining bias in parental reports: Both the accuracy and distortion hypothesis predict the reporting of more problematic behavior in children by parents with psychiatric symptoms (Richters, 1992). To establish whether these reports reflect real child problems or distorted parental perceptions, criterion ratings are needed, "which are

themselves validated, independent of the influence of maternal depression and are based on sampling both situations and behaviors comparable to those sampled by mothers' ratings" (Richters, 1992, p.487). The sampling requirement is especially hard to meet in field research. Most studies reviewed by Richters relied on children, observers or teachers as criterion informants, rating child behavior problems in different situations. Therefore, differences between informants could easily be ascribed to well known differences in child behavior across situations (Achenbach, McConaughy, & Howell, 1987; Stanger & Lewis, 1993). In one of the very few studies conducting laboratory research on this subject, Youngstrom et al. (1999) compared mothers' ratings of videotaped behavior samples of their children with the ratings of independent observers, judging positive and negative child behaviors and emotions. Multiple regression techniques indicated correlations between maternal ratings and lab ratings ranging from .32 to .41 for various child outcomes. After removing variance shared with lab ratings from mothers' child behavior ratings, maternal dysphoria explained 2.3% to 20% of the residual variance. Interestingly, correlations between different informants reported in this laboratory study closely parallel those generally found in field research (cf. Achenbach et al.,1987), and variance explained by maternal distress under controlled laboratory conditions is found to be equal to the amount of depression-related variance reported in field studies (Fergusson et al., 1993).

In summary, despite problems of sampling and criterion validity, there is growing evidence that parental distress or psychopathology is significantly related to emotional and behavior problems in their children *and* may cause small to moderate parental reporting distortions. Our study aims at establishing the amount of accuracy and distortion of maternal child reports in a clinical sample of 6- to 12-year-old boys. The effect of a broad range of maternal psychopathological symptoms on mothers' ratings of different types of child behavior problems is examined using teachers and group care workers together as criterion informants. Thus, type of maternal psychopathology and type of child behavior problems vary, but sex and age of the child are held constant. A unique feature in this study is the use of multiple criterion informants. Our clinical setting allows for the use of both teachers and group care workers as skilled observers with access to a great variety of child behaviors. Other than teachers, group care workers observe children in a situation that resembles home more than the classroom. Taken together, the reports of teachers and group care workers cover a large sample of child behavior problems, thereby providing a more robust criterion against which to evaluate maternal reports than most field studies have done so far. Multiple regression techniques are used to quantify both the variance shared by mothers and criterion raters, and the residual variance accounted for by maternal psychopathology.

METHOD

Subjects

Subjects in this study were 6- to 12-year-old children receiving treatment in the residential and day care units of a child welfare institution in the Netherlands (*Paedological Institute* in Nijmegen). All children having regular contact with their biological mother during the period of treatment were selected for the study. Their mothers were asked to participate and 89 (55%) of these responded positively. Due to missing data from one or more informants, the sample was reduced to 83 subjects, 68 boys and 15 girls. To create a homogeneous sample regarding to the sex of the subjects, only boys were selected for the present study. In this final sample, 41% of the boys were in residential treatment and 59% in day treatment. Expert child psychiatrists and psychologists in a multidisciplinary assessment center made child diagnoses, which led to placement in the clinic. None of the children met the criteria for mental retardation. Mean age of the 68 boys in our sample was 8.7 years, mean age of mothers 38.9. Mothers' occupational status was categorized using the Social Demographic Inventory (SDI; Van Westerlaak, Kropman, & Collaris, 1975). Sixty percent of the mothers were housewives, 13% held jobs requiring only high school education, 24% had an occupation requiring some kind of college degree, 2% were engaged in an academic career. All children received special education as an integral part of the treatment at a school associated with the Paedological Institute. Teachers and group care workers participating in the study had known the boys for at least six months at the moment of their assessment of the child's behavior.

Measures

CBCL and TRF

Mothers and group care workers rated children's internalizing and externalizing behavior problems using the Dutch translation of the Child Behavior Checklist (CBCL; Achenbach, 1991; Verhulst, Van der Ende, & Koot, 1996). By reporting for each of 118 specific behaviors on a three-point scale, informants give a detailed picture of a child's behavioral problems. These behavioral items can be grouped around eight factors or narrow-band syndromes, which in their turn can be grouped under two broad-band syndromes, designated as Internalizing and Externalizing. The Internalizing and Externalizing scores were used in our analysis. Teachers completed the parallel version of the CBCL, the Teacher's Report Form (TRF; Achenbach, 1991), also translated into Dutch (Verhulst, Van der Ende, & Koot, 1997). The CBCL and TRF Internalizing and Externalizing scales are intended to reflect common elements in the child's behavior and consist largely of identical items. Yet,

some items are unique to one or the other version because of differences in potential access of parents versus teachers to particular kinds of problem behaviors (for instance Disobedient at home versus Disrupts class). In order to sample comparable behaviors for mothers and teachers (Richters, 1992), only those items were selected that are common to both CBCL and TRF. Raw scores for Internalizing and Externalizing scales were computed from these common items and used in the analyses.

The factor analytic structure of the CBCL filled out by group care workers was recently studied by Albrecht, Veerman, Damen, and Kroes (2001). Confirmatory factor analysis showed that the original CBCL factor model based on parental judgments also fits for the judgment of group care workers. In order to sample comparable behaviors for mothers and group care workers, only those items were selected that are common to both CBCL and TRF. Raw scores for Internalizing and Externalizing scales were computed from these common items and used in the analyses.

Again, raw scores for the Internalizing and Externalizing scales based on items common to both CBCL and TRF were used in the analyses.

SCL-90

To assess maternal psychopathology mothers completed the Dutch version of the (Revised) Symptom Checklist (SCL-90; Derogatis, 1977; Derogatis, Lipman, & Covi, 1973; Arrindell & Ettema, 1986). This instrument, originating from the Hopkins Symptom Checklist (HSCL), was designed to measure a broad range of psychopathology and has proven to be sensitive in discriminating between different levels of pathology in psychiatric outpatients as well as in non-clinical samples (Arrindell & Ettema, 1986; Derogatis, Lipman , & Covi, 1983). The Dutch version of the SCL-90 was constructed following the original scale and modified to accommodate to the Dutch population. Factor analytic studies in clinical and non-clinical Dutch samples established eight dimensions mostly resembling the original nine dimensions derived by Derogatis (1977). The Dutch scale dimensions are Agoraphobia, Anxiety, Depression, Somatic complaints, Thought insufficiency (labeled Obsessive-compulsive behavior by Derogatis), Interpersonal sensitivity, Hostility and Sleeping problems (the latter not corresponding to one of the original dimensions). A general, ninth dimension was added, comprising the sum total of all 90 items. This dimension is labeled Neuroticism. Both reliability and validity of the Dutch version proved to be satisfactory in large sample studies (Arrindell & Ettema, 1986). In the present study, general population norms for women were applied.

Data analysis

First, descriptive data are presented on children's behavior problems as reported by all three informants. Statistical tests of differences between children of mothers with and without

psychopathology were performed, using the 65th percentile of the SCL-90 — indicating 'above average' levels of psychopathology — as the cut-off point. Besides the mean scores, effect sizes (ES; Cohen, 1988) of the differences between ratings for children of mothers with and without psychopathology are given. According to Cohen, d = .20 is to be considered a small, d = .50 a medium, and d = .80 a large ES for the difference between independent means.

Next, stepwise multiple regression analysis was used to estimate the amount of variance shared by mother and criterion informant reports, and the effect of mothers' psychopathological symptoms on their reporting of child behavior problems. Separate analyses were conducted for internalizing and externalizing child behavior problems. Mothers' CBCL-ratings of internalizing or externalizing child behaviors served as the dependent measure in all regressions. Group care workers' CBCL-ratings and teachers' TRF-ratings (using only common scale items) of internalizing or externalizing child behaviors entered the regression model together as independent variables in the first step. Subsequently, one of the SCL-90 psychopathology dimensions was included as the only independent variable in the second step. Separate analyses were conducted for each of the nine dimensions of psychopathology. High correlations between maternal child ratings and criterion informant ratings indicate agreement between informants on the existence of child behavior problems. If, after regressing criterion informant ratings on maternal child ratings, mothers' psychopathological symptoms predict a substantial amount of residual variance, this would suggest over-reporting by mothers with high levels of psychopathological symptoms.

RESULTS

Descriptive data on internalizing child behavior problem scores by mothers, group care workers and teachers, for boys of mothers with and without different forms of psychopathology are presented in Table 1. Group means are shown, as well as effect sizes of group differences and statistical significance of independent sample t-tests. Externalizing behavior problem scores are presented in Table 2.

Table 1.

Mothers', Group Care Workers' and Teachers' Internalizing Behavior of Boys in Relation to Mother's Psychopathology. Means and Effect Sizes of Differences between the Means.

		Mother		Group care worker		Teacher	
SCL-90 symptoms	*N*	*M*	*d*	*M*	*d*	*M*	*d*
Neuroticism							
Present	34	16.50	.71**	10.62	.24	7.97	-.27
Not present	34	10.76		9.15		9.74	
Anxiety							
Present	31	16.71	.70**	10.42	.17	8.19	-.18
Not present	37	11.05		9.43		9.41	
Phobia							
Present	37	16.05	.65**	10.05	.06	8.78	-.02
Not present	31	10.74		9.68		8.94	
Depression							
Present	39	16.08	.71**	10.56	.27	8.72	-.05
Not present	29	10.34		8.97		9.03	
Somatization							
Present	29	16.34	.57*	11.07	.35	7.52	-.35
Not present	39	11.62		9.00		9.85	
Thought insufficiency							
Present	29	17.21	.78**	11.21	.39	8.72	-.03
Not present	39	10.97		8.90		8.95	
Interpersonal sensitivity							
Present	33	15.79	.50*	10.45	.18	9.09	.07
Not present	35	11.60		9.34		8.63	
Hostility							
Present	44	14.98	.45	10.14	.12	9.05	.08
Not present	24	11.17		9.42		8.50	
Sleeping problems							
Present	31	17.52	.92**	11.16	.40	8.65	-.06
Not present	37	10.38		8.81		9.03	

Note. Differences between the means were tested with the *t* test for Independent Samples. * $p < .05$, ** $p < .01$.

Results indicate that mothers with above average levels of psychopathological symptoms generally reported more internalizing and externalizing behavior problems in their sons than mothers with average or low levels of psychopathology. Differences between reports of internalizing behavior problems reached statistical significance (p <.05) for all dimensions of psychopathology, except for hostility.

Table 2.

Mothers', Group Care Workers' and Teachers' Externalizing Behavior of Boys in Relation to Mother's Psychopathology. Means and Effect Sizes of Differences between the Means.

		Mother		Group care worker		Teacher	
SCL-90 symptoms	*N*	*M*	*d*	*M*	*d*	*M*	*d*
Neuroticism							
Present	34	21.50	.47	17.06	.33	14.09	.40
Not present	34	17.12		13.59		9.88	
Anxiety							
Present	31	21.77	.48	16.77	.26	14.32	.41
Not present	37	17.24		14.11		10.03	
Phobia							
Present	37	21.59	.54*	16.03	.15	13.49	.31
Not present	31	16.58		14.48		10.19	
Depression							
Present	39	22.23	.76**	17.18	.42	13.44	.32
Not present	29	15.38		12.83		10.03	
Somatization							
Present	29	23.62	.84**	17.17	.31	13.97	.32
Not present	39	16.10		13.95		10.51	
Thought insufficiency							
Present	29	21.93	.49	16.55	.20	13.86	.31
Not present	39	17.36		14.41		10.59	
Interpersonal sensitivity							
Present	33	22.18	.60*	17.88	.49	14.21	.41
Not present	35	16.60		12.91		9.89	
Hostility							
Present	44	21.16	.56*	16.11	.21	12.91	.24
Not present	24	15.92		13.88		10.29	
Sleeping problems							
Present	31	20.84	.29	14.58	.13	12.03	.01
Not present	37	18.03		15.95		11.95	

Note. Differences between the means were tested with the *t* test for Independent Samples. * $p < .05$, ** $p < .01$.

For externalizing behavior problems, differences in ratings between mothers with and without psychopathology reached statistical significance for only five out of nine dimensions of psychopathology. Group care workers and teachers, on the other hand, did not observe any statistically significant differences in behavior problems between children of mothers with and without psychopathology. When comparing the child behavior problem scores of all three

informants for children of mothers without psychopathology, it can be seen that mothers generally reported higher levels of externalizing behaviors than the other informants, whereas all informants observed comparable levels of internalizing child behavior ratings.

Table 3.

Regressions of Maternal Internalizing Child Behavior Ratings on Maternal Psychopathological Symptoms after Regression on Criterion Child Behavior Ratings. N = 68.

Predictors	*R*	R^2	ΔR^2	*Δp*
Step 1.				
Criterion informants	.448	.201		.001
Step 2.				
Neuroticism	.613	.376	.175	.000
Anxiety	.618	.382	.181	.000
Phobia	.554	.307	.106	.003
Depression	.561	.315	.114	.002
Somatization	.510	.260	.059	.027
Thought insufficiency	.634	.402	.201	.000
Interpersonal sensitivity	.565	.319	.118	.002
Hostility	.581	.338	.137	.001
Sleeping problems	.571	.326	.125	.001

Note. A separate analysis is conducted for each of the SCL-90 psychopathological symptoms entering in the second step, after the informant block.

Stepwise multiple regressions were conducted to analyze the relative contributions of criterion informant ratings and psychopathological symptoms to the prediction of maternal child ratings. As shown in Table 3 and Table 4, ratings of both criterion informants — entered together — explained a significant amount of variance in maternal child behavior ratings. Cross-informant correlation was higher for externalizing behavior problems (R = .58) than for internalizing behavior problems (R = .45). For internalizing behavior, mothers' psychopathological symptoms — entered in the second step in separate analyses — accounted for an additional 6% to 20% of variance in their child ratings (14% on average).

All dimensions of psychopathology made a statistically significant contribution to the prediction of mothers' ratings of their children's internalizing behavior problems. For externalizing behavior, the additional variance in mothers' ratings explained by maternal

psychopathology appeared to be much smaller, ranging from 2% to 7% (5% on average). Only six out of nine dimensions of psychopathology made a statistically significant improvement in the prediction of mothers' externalizing child behavior ratings.

Table 4.

Regressions of Maternal Externalizing Child Behavior Ratings on Maternal Psychopathological Symptoms after Regression on Criterion Child Behavior Ratings. N = 68.

Predictors	*R*	R^2	ΔR^2	*Δp*
Step 1.				
Criterion informants	.579	.336		.000
Step 2.				
Neuroticism	.631	.398	.062	.013
Anxiety	.630	.400	.061	.014
Phobia	.600	.357	.021	.150
Depression	.602	.362	.026	.107
Somatization	.635	.403	.067	.009
Thought insufficiency	.615	.378	.042	.040
Interpersonal sensitivity	.628	.395	.059	.015
Hostility	.627	.394	.058	.016
Sleeping problems	.594	.353	.017	.196

Note. A separate analysis is conducted for each of the SCL-90 psychopathological symptoms entering in the second step, after the informant block.

DISCUSSION

The aim of our study was to establish the amount of accuracy or distortion in maternal ratings of their child's behavior problems due to mother's own psychopathological symptoms. The 68 subjects in our study were 6- to 12-year old, clinic-referred boys. The effect of a broad range of maternal psychopathological symptoms on mothers' ratings of different types of child behavior problems was examined, using group care workers and teachers as criterion informants. Multiple regression techniques were used to quantify the amount of variance shared by mothers and criterion raters — reflecting accuracy, as well as the amount of residual variance accounted for by maternal psychopathology — reflecting distortion.

Maternal psychopathology explained small to moderate (2% to 20%) variance in mothers' child behavior ratings, after partialing out variance shared with criterion ratings. The amount of residual variance explained by maternal psychopathology parallels the amounts of variance due to maternal dysphoria or depression found in the lab (2.3% to 20%; Youngstrom et al., 1999) as well as in field research (1.7% to 16%; Fergusson et al., 1993). However, the amount of additional variance explained by maternal psychopathology varied largely according to the type of child behavior problems. For externalizing behavior problems, maternal psychopathology added little or nothing to the prediction of mothers' child ratings, when controlling for criterion information. For internalizing behavior problems, on the other hand, the amount of additional variance in mothers' reports explained by various dimensions of maternal psychopathology was substantial and almost as large as the amount of variance shared with criterion informants. In addition, the amount of variance explained by maternal psychopathology varied considerably according to the type of psychopathological symptoms. For example, maternal hostility or interpersonal sensitivity seem to produce small distortions in mothers' reports of externalizing child behavior problems, whereas maternal phobia or depression do not produce this kind of distortion. This underscores the differential role of the type of psychopathological symptoms.

The above findings are consistent with the general view that various forms of maternal psychopathology cause small to moderate distortions in mothers' reports of child behavior problems. At the same time, our results suggest a more detailed specification of the impact of maternal psychopathology on child behavior ratings. That is, the distorting effect of maternal psychopathology seems to be primarily related to the reporting of internalizing child behavior problems. Further, different forms of maternal psychopathology seem to play a differential role in producing distortion of maternal child behavior ratings. And third, type of maternal psychopathology and type of reported child behavior seem to somehow interact to produce distortion, or not.

Several explanations have been suggested to account for the phenomenon that maternal psychology may lead to bias or distortion in the perception of child behavior problems (Richters, 1992). Of these, the projection hypothesis (cf. Moretti et al., 1985) seems most consistent with our results. This hypothesis assumes that mothers project symptoms of their own psychological states in their children. As Briggs-Gowan et al. (1996, p. 751) suggest, this "may be most likely with internalizing symptoms that may be ambiguous or require a high degree of inference". Another potential explanation referring to the ambiguity of behavioral stimuli is provided by the social attribution theory (cf. Dodge, 1986). This theory asserts that ambiguous environmental stimuli are more liable to inference and distortion of social perception than more obvious stimuli. In line with this view, research has shown that both anxious and aggressive parents and children display cognitive biases in reaction to ambiguous stimuli (Barrett, Rapee, Dadds, & Ryan, 1996; Strassberg, 1995, 1997). In accordance with the social attribution theory, our results indicate that internalizing — or

more ambiguous — behavior problems seem more liable to perceptual distortions induced by various kinds of psychopathological symptoms. On the other hand, with externalizing behavior problems, less distortion caused by maternal psychopathology can be expected because externalizing problems are more readily observable.

So far, additional variance in mothers' reports associated with maternal psychopathological symptoms was interpreted as reflecting bias or distortion in mothers' perception. However , an alternative interpretation of our data cannot be ruled out, i.e. that children of mothers with high levels of psychopathological symptoms actually display more internalizing problems at home, which are either not fully displayed or not accurately observed in the clinic or school. Generally, internalizing child behaviors appear to be less stable across situations than externalizing behaviors, as several multitrait-multimethod studies have suggested (Kazdin, Esveldt-Dawson, Unis, & Rancurello, 1983; Saylor, Finch, Baskin, Furey, & Kelly, 1984; Stanger & Lewis, 1993). Thus, internalizing child behavior problems associated with maternal psychopathological symptoms, displayed at home, might not be manifested in other situations, or they might be manifested in a different way under different circumstances. Finally, It should be stressed that our findings apply to a clinical sample of 6- to 12-year old boys, and therefore cannot be generalized easily to other sex or age groups, or community samples. Research under more controlled conditions is needed to further investigate the complex relationship between maternal psychopathological symptoms and the manifestation and perception of internalizing child behavior problems.

Chapter 4

The Impact of the Big Five Personality Traits on Reports of Child Behavior Problems by Different Informants

Gert Kroes, Jan W. Veerman, and Eric E. J. De Bruyn

The present study compared ratings of a standardized sample of child behavior problems across informants and examined the effects of informant personality traits on child behavior ratings by mothers, teachers and group-care workers. Subjects were 55 clinic-referred children, aged 6-12. All informants watched and rated the same 17-minute videotaped behavior sample of a familiar target child. Independent trained observers rated the same videotapes to provide criterion ratings. Informants' personality traits were assessed using the NEO Five Factor Personality Inventory. Results showed that mothers reported fewer behavior problems than did the professionals, that the informants who were familiar with the child reported more behavior problems than did the independent observers, and that higher levels of informant neuroticism were related to higher ratings of child behavior problems in the case of the professionals, but not in the case of the mothers. In addition, group-care workers who were less extraverted and open were likely to report more child behavior problems than group-care workers with normal levels of extraversion and openness. Finally, no relations were found between agreeableness or conscientiousness and ratings of child behavior. Findings suggest that professionals who work with children are not immune to distortions based on their own personality.

INTRODUCTION

Ratings of child behavior problems vary across informants and situations. Invariably, only modest agreement is found among different informants' ratings of a child's functioning (Achenbach, McConaughy, & Howell, 1987; Stanger & Lewis, 1993). This phenomenon constitutes a major obstacle for both researchers and clinicians trying to assess child behavior and emotional problems accurately. Various reasons for the disagreement have been suggested: general mechanisms such as response set, social desirability, or a willingness to report negative behaviors on the part of the informant (Youngstrom, Loeber, & Stouthamer-Loeber, 2000); informant personality factors such as depression and anxiety (Briggs-Gowan, Carter, & Schwab-Stone,1996); and interactional factors such as the parent-child relationship (Treutler & Epkins, 2003). In addition, child behaviors can vary across the situations of home, school, neighborhood, and clinic (Kolko & Kazdin, 1993). Given that different informants observe children in different situations, situational factors may also play an important role in the creation of disagreement among informants. In the present study, the impact of personality factors of the informant on reports of child behavior problems was examined while controlling for situational differences.

The effects of informant personality factors, particularly maternal depression,on reports of child behavior problems have frequently been examined (cf. Chilcoat & Breslau, 1997; Kroes, Veerman, & De Bruyn, 2003; Youngstrom, Izard, & Ackerman, 1999). In many studies, higher levels of child behavior problems reported by mothers were found to be associated with higher levels of depressive symptoms in the mothers themselves. The question is how to interpret these findings: Do the reports of the mothers with higher levels of depression actually reflect increased levels of behavior problems in their children, or does maternal depression distort their perceptions of child behavior? Richters (1992) described these two different perspectives on the perception of child behavior as the *accuracy model* and the *distortion model*. In many studies, maternal psychopathology has indeed been argued to cause distortion (e.g., Jensen, Traylor, Xenakis, & Davis, 1988). In a few other studies, however, no evidence for the distortion hypothesis has been found (e.g., Conrad & Hammen, 1989) or the bias in the mothers' reports of child behavior problems has been minimal, and thus of little or no clinical significance (Sawyer, Streiner, & Baghurst, 1998). The most recent studies report mixed results, which suggests that both models may apply: maternal psychopathology may be associated with increased behavior problems among children, *and* mothers with psychopathological symptoms may overreport the incidence and/or severity of such problems (e.g., Chilcoat & Breslau, 1997; Najman et al., 2000).

Most of the aforementioned studies have been hampered by methodological problems. As Richters (1992) has pointed out, *both* the accuracy hypothesis and the distortion hypothesis predict reports of more problematic child behavior by parents with psychiatric symptoms. To establish whether these reports reflect real child problems or distorted parental perceptions, there is a need for criterion ratings which are themselves "validated, independent of the influence of maternal depression" and based on the sampling of both "situations and behaviors comparable to those sampled by mothers' ratings" (Richters, p. 487). In all studies reviewed by Richters, situations and behaviors differed across informants. Therefore, differences between mothers' and criterion ratings can easily be ascribed to well-known differences in child behavior across situations (Achenbach et al., 1987). As a consequence, the studies available to date fail to provide conclusive evidence for the distorting impact of informant psychopathology. The most important design requirement for testing the distortion hypothesis is thus standardization of the child behavior sampling frame for both the mothers and criterion informants. When this condition has been met, evidence of the following will then be needed to provide support for the distortion hypothesis: "(a) depression-related disagreements between mothers and criterion raters and (b) the superiority (e.g., accuracy) of criterion ratings over mothers' ratings" (Richters, p. 487). Finally, Richters has also suggested that field studies should be complemented with laboratory studies involving, for example, videotaped samples of child behavior, thereby allowing greater control over the child behaviors being observed.

Only a few studies have been intentionally designed to meet Richters' (1992) methodological requirements to date. In a laboratory study by Youngstrom et al. (1999), mothers' ratings of videotaped behavior samples from their children were compared to the ratings of independent observers. Multiple regression analyses revealed correlations between maternal ratings and observer ratings ranging from .32 to .41 for various types of child behaviors. After removal of the variance shared by the maternal and observer ratings, maternal dysphoria was then found to explain 2.3% to 20% of the residual variance. Interestingly, the correlations found between the different informants in this laboratory study closely parallel the correlations typically found in field research (cf. Achenbach et al., 1987), and the amount of variance explained by maternal distress under well-controlled laboratory conditions equals the amount of variance explained by depression in field studies (Fergusson, Lynskey, & Horwood, 1993).

In two related studies, Johnston and Short (1993) examined the relations of depressive symptomatology to perceptions of child behavior for both mothers and female undergraduate students. More specifically, the adult females were asked to complete measures of depressive symptomatology and rate the videotaped behaviors of child actors. For the mothers, depressive symptoms were correlated with more negative perceptions of internalizing child behaviors and of prosocial behaviors. Female students with more depressive symptoms had

less positive perceptions of prosocial behaviors and gave more negative global ratings for the externalizing behaviors.

The results of the studies by Johnston and Short (1993) suggest that depressive symptomatology may negatively influence adult perceptions of child behavior, and that perceptual distortion is not limited to depressed mothers judging their own problem children but also applies to other adults with depressive symptoms, judging unfamiliar children. This stresses the need, as already noted by Richters (1992), to examine the influence of depression on adults working professionally with children, which may include teachers and child-care workers. It also raises the question of whether familiarity with the child moderates distortion, as is suggested by several studies (Kendziora and O'Leary, 1998; Lorber, O'Leary, & Kendziora, 2003; Snarr, Strassberg, & Slep, 2003). For example, Kendziora and O'Leary (1998) compared mothers' ratings of videotapes of their own and unfamiliar children's behavior to independent observers' ratings. They found that mothers reported fewer negative behaviors for their own children than did the observer, and also that mothers evaluated their own children's behavior as less negative than unfamiliar children's behavior.

Although the focus of considerable research has been on perceptual distortion elicited by depression, the influences of other informant personality factors have also been studied. Anxiety or substance use (Chilcoat & Breslau, 1997), dysphoria or distress (Briggs-Gowan et al, 1996; Youngstrom et al., 1999), parental aggression (Strassberg, 1995, 1997), as well as various other psychopathological symptoms (Jensen et al., 1988; Kroes et al., 2003) have been found to be associated with perceptual distortion in mothers. These studies suggest that research on perceptual distortion should be expanded to a more general examination of the influence of various personality traits on reports of child behavior problems.

The Five-Factor Model of personality (McCrae & Costa, 1999) provides the most comprehensive and widely accepted set of personality traits. We decided to examine the association between perceptual distortion and personality traits for each of the Big Five factors (Neuroticism, Extraversion, Openness, Agreeableness and Conscientiousness). More specifically, we studied the influence of these traits on the ratings of child behavior problems in a multi-informant design using videotaped samples of child behavior. With the use of videotaped behavior samples, a standardized sampling frame was created for both the mothers and professionals (i.e. teachers and other child-care workers), rating familiar children. In addition, trained observers who were unfamiliar with the children watched and rated the same videotapes to provide a completely independent criterion for comparison to the ratings provided by the mothers and professionals. Finally, both the mothers and child-care professionals completed a standardized measure of the Big Five personality traits.

The design of this study allows for the comparison of mothers' and professionals' ratings of identical samples of a given child's behavior, while controlling for familiarity with the child. Moreover, the effect of familiarity can be tested by comparing mothers' and professionals' ratings to the ratings by independent observers who were not previously

familiar with the child. A unique feature of the study is that the design allows the distortion hypothesis to be tested for both mothers and professionals while meeting the research requirements stipulated by Richters (1992).

We addressed the following three research questions. First, we tested for differences in child behavior ratings of mothers and professionals under controlled conditions. In view of informant differences generally found in both natural and laboratory settings (Kroes et al., 2003; Youngstrom et al., 1999), we expected that mothers would report more behavior problems than would professionals. Second, we tested the effect of familiarity with the child. In line with past research, we hypothesized that mothers rating their own children would report fewer behavior problems than would independent observers who had not previously known the child. This might be interpreted as reflecting a positive bias in the mothers (Kendziora & O'Leary, 1998). It is unknown whether such a positive bias can be generalized to professionals rating familiar children; therefore, the comparison between the professionals' ratings and the independent observers' ratings was exploratory.

Third, we tested whether any of the Big Five personality traits in mothers or professionals appear to produce a distortion in the child behavior ratings provided by these informants (i.e. ratings that clearly deviate from the independent criterion ratings). Many studies discussed so far demonstrate the distorting effect of depression and anxiety, which are actually facets of Neuroticism in the Big Five model (Costa & McCrae, 1992). Therefore, we expected that higher levels of Neuroticism would be related to higher ratings of child behavior problems. In a general review of personality judgment research, Funder (1999) notes that the effect of other aspects of the judges' personality on their ability to observe the behavior of others accurately has not been systematically studied. Besides Neuroticism, Extraversion and Openness are the traits that most consistently seem to be associated with judgment accuracy (Funder, 1999). People who are outgoing and open to experiences might be more likely to accept individual differences or have less rigid expectations of what constitutes normative behavior. We therefore expected that lower levels of Extraversion or Openness would be related to higher ratings of child behavior problems. As far as we are aware, no systematic associations between Agreeableness or Conscientiousness and perceptual distortion are suggested in the literature.

METHOD

Participants

The subjects were 55 elementary school-aged children (43 boys and 12 girls) referred to a Dutch clinic for the treatment of youths with emotional, behavioral, and learning problems. During two consecutive years of admission, all children attending the primary

school of the clinic were recruited for the study and their mothers were asked to participate. From an initial pool of 63 children whose mothers had agreed to take part in the study, eight cases were omitted due to: incomplete data (3), premature termination of treatment (2), maternal illness (2), or withdrawal of consent to participate by the mother (1). The mean age of the children in the remaining sample of 55 children was 8.9 years (*SD* = 1.9, range of 6 to 13 years). All children were diagnosed with psychiatric disorders, such as Attention Deficit Hyperactivity Disorder, Pervasive Developmental Disorder, Oppositional Defiant Disorder or Conduct Disorder, as defined by the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (American Psychiatric Association, 1994). None of the children met the criteria for mental retardation. The mean age of the mothers was 38.5 years (*SD* = 4.8, range of 25 to 52 years). The occupational status of the mothers was categorized using the Dutch Social Demographic Inventory (SDI; Van Westerlaak, Kropman, & Collaris, 1975); on a 6-point scale ranging from (1) unskilled labor to (6) academic career, the median score was 3.

The 55 children participating in the study all received either residential (47%) or day treatment (53%), with special education constituting an integral part of the treatment. Their teachers were asked to participate in the study in addition to their group-care workers. Seventy-eight percent of the teachers who were asked to participate responded positively, as well as 80% of the group-care workers. Negative responses were due to job change, illness, or refusal to participate. The final sample included 12 teachers, rating 42 children (*M* = 3.5, *SD* = 2.54, range 1 to 10 children per teacher) and 26 group-care workers, rating 44 children (*M* = 1.69, *SD* = 1.05, range 1 to 5 children per group-care worker). The mean age of the teachers was 39.2 years (*SD* = 7.4), and 58% of the teachers were female. The mean age of the group-care workers was 33.0 years (*SD* = 6.7), and 81% were female. All of the teachers and group-care workers had a college degree in education. The professional experience of the teachers ranged from 4 to 21 years (*M* = 8.3, *SD* = 5.4); the professional experience of the group-care workers ranged from 1 to 20 years (*M* = 6.8, *SD* = 5.1). Although the group-care workers in the residential treatment group spent somewhat more time with the children, the roles of group-care workers in residential and day treatment were essentially the same. All of the professionals participating in the study had known the target child for at least six months.

Measures

NEO Five Factor Inventory (NEO-FFI)

Informant personality traits were assessed using the NEO-FFI (Costa & McCrae, 1992; Dutch version by Hoekstra, Ormel, & De Fruyt, 1996). The NEO-FFI is the short form of the NEO-PI-R (Hoekstra et al., 1996), which is an authorized translation of the Revised NEO Personality Inventory by Costa and McCrae (1992). The NEO is a widely used and

well-validated self-report measure of the Big Five personality traits of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. The NEO-FFI consists of 60 items to be rated along a 5-point scale. It takes 10-20 minutes to complete. The norms for the Dutch NEO-FFI were obtained from a general population sample (N = 2415). The NEO-FFI raw factor scale scores are transformed into stanine scale scores (with a mean of 5 and a standard deviation of 2) using Dutch population norms.

Direct Observation Form (DOF)

The DOF (Achenbach, 1986; Dutch version by Ferdinand & Verhulst, 1998) was designed to provide an efficient observational instrument that covers a wide range of behavior problems. It can be used by teacher aides and research assistants in applied settings such as the home and the classroom. The DOF consists of 96 items, such as "Doesn't sit still," "Shows off," "Is nervous or tense." Each item is rated along a 0-1-2-3 response scale, with 0 indicating no observed occurrence of the behavior, 1 indicating a slight or ambiguous occurrence, 2 indicating definite occurrence with a mild to moderate intensity and a duration of less than 3 minutes, and 3 indicating definite occurrence with severe intensity or a duration of 3 or more minutes. In our study, we used the Total Problems score, which is derived from the sum of all 96 items.

The DOF is normally completed after 10 minutes of child observation. For the purposes of the present study, which included the rating of videotapes, an alternative assessment procedure was followed. Rather than complete the DOF form, the informants were asked to sort a deck of 96 cards with the DOF items on them. In the first step of the rating procedure directly following the viewing of the videotape, the mothers, teachers, group-care workers, and independent observers were asked to select those DOF items that they definitely did not observe on the videotape. The cards containing these items were then removed from the deck and assigned a score of 0 by the interviewers. In the second step of the rating procedure, the informants were asked to sort the remaining cards into three piles reflecting the DOF scale anchors (i.e. rating scores of 1, 2 and 3 respectively). Written descriptions of the scale anchors were provided to facilitate sorting. The interviewers subsequently transferred the results of the sorting procedure onto the DOF form. This two-step administration procedure was designed to attain scoring consistency across the different informants by asking them to make deliberate choices about each item and providing them with written descriptions of the anchor points throughout the decision-making process.

Procedure

Videotapes of the children were made in the play group at the treatment center using a script that was tested in a previous study (Muntinga, 1999). The video script was designed to both standardize the behavior sample and to elicit a range of naturalistic child behaviors. Two consecutive but different episodes were recorded during the regular afternoon group play session for each of the target children. First, the target child was instructed to invite three other children to play a board game with him or her in front of the camera, and the invitation process was then recorded for five minutes. Second, the children's play was subsequently recorded for 12 minutes. The play was supervised by an assistant group-care worker who was instructed to intervene only when deemed necessary to maintain order in the group. The two video fragments were then combined into a 17-minute videotaped behavior sample for each of the target children.

The mothers, teachers, and group-care workers were asked to watch the videotapes individually and assess the child behavior problems observed on the videotapes during an interview session at the clinic. The mothers signed consents forms and received a gift worth 15 euros after completing the interview. Teachers and group-care workers received 15 euros for the rating of a videotape. The individual interviews were conducted by trained research assistants (N = 5), who were unfamiliar with the target children. The research assistant interviewing a particular informant was also unaware of the ratings provided by the other informants for the child in question. During the interview sessions, the mothers, teachers and group-care workers first completed the NEO-FFI. Immediately thereafter, they viewed the videotape and were asked to assess the child behavior problems that they may have observed on the tape using the DOF.

To provide the independent criterion measures, two observers individually viewed the videotapes and completed the DOF for each target child. The independent observers were undergraduate psychology students with previous training in the domain of child observation and assessment. The independent observers received additional training to use the DOF, which included the rating of pilot videotapes simultaneously observed by an experienced observer. Both the mothers, professionals and independent observers were instructed to rate only those behaviors that they actually observed on the videotape. In addition, the observers were instructed to review the videotape as many times as they felt necessary. Observer training was continued until inter-rater Kappa's across sets of DOF items reached at least .80 for a set of five pilot videotapes. The Pearson correlation between the independent observers' individual ratings of all videotapes used in the study (N = 55) was .85. After individually rating the videotapes, the observers discussed remaining scoring discrepancies until consensus was reached. In all subsequent statistical analyses, the consensus score was used as the criterion measure.

Data analysis

Informant differences were tested by conducting a one-way repeated-measures ANOVA on the DOF video ratings, with the four categories of informant constituting the independent variable. We tested the following three planned contrasts: (1) mothers vs. teachers and group-care workers, (2) mothers vs. independent observers, and (3) teachers and group-care workers vs. independent observers. Only those children rated by all four informants (N = 34) were used in this analysis. Missing data analyses showed that the missing cases did not significantly differ (p >.05) from the complete sample regarding maternal DOF ratings, independent observer DOF ratings, sex and age of the children, or type of treatment.

To test for the effect of personality traits on child behavior ratings, partial correlations were computed between each of the informant's NEO personality traits and the respective informant's DOF ratings, while controlling for the independent observers' criterion ratings. In other words, we tested whether any of a given informants' personality traits would explain variance in their child behavior ratings, after partialing out variance in their child behavior ratings shared with the independent observers. For Neuroticism, Extraversion, and Openness, probability tests were one-tailed because of the directional nature of the hypotheses.

In the case of teachers and group-care workers, dependence among their child behavior observations complicated the analysis. While the mothers in our sample provided independent observations of their children, the teachers and group-care workers often rated more than one child. Dependence among their observations might create spurious correlations between personality traits and child behavior ratings when testing for personality-related distortions (Cohen, Cohen, West, & Aiken, 2003; Stevens, 2002). We addressed this issue by aggregating child behavior ratings at the group level, obtaining a mean child behavior rating for each teacher and group-care worker. Teachers and group-care workers were then treated as the unit of analysis, instead of the total number of children rated by these informants. Technically, multilevel analysis would be a more appropriate approach to account for the dependency of observations in our sample, but there is not enough nested data to justify the use of such a substantially more complicated method.

RESULTS

Descriptive Statistics

As can be seen in Table 1, the mothers in our sample displayed average levels of Neuroticism, Extraversion, Openness and Conscientiousness, but higher levels of

Agreeableness, as compared to the general Dutch population (one-sample *t* test, $p < .01$). Both the teachers and group-care workers displayed significantly lower levels of Neuroticism, as compared to the Dutch population, and higher levels of Extraversion and Openness; group-care workers also showed higher levels of Agreeableness (all *p*'s < .05).

Table 1.

NEO Personality Traits for Different Informants (Stanine scale scores).

		Neuroticism		Extraversion		Openness		Agreeable-ness		Conscien-tiousness	
Informant	*N*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Mother	55	5.11	1.89	5.24	2.01	5.05	1.89	5.71	1.88	5.22	2.23
Teacher	12	3.17	1.53	6.25	1.66	7.42	1.00	5.33	1.23	4.75	1.71
Group-care worker	26	4.08	1.47	6.35	1.72	7.08	1.13	5.92	1.70	5.27	2.54

Given the sampling differences, the DOF Total Problems scores obtained in our study (Table 2) can only tentatively be compared with the DOF scores reported in the Manual for the ASEBA school-age forms & profiles (Achenbach & Rescorla, 2001). The mean DOF scores provided in the Manual were obtained by averaging over 3 to 6 10-minute observations of the classroom behaviors of elementary school-aged boys and girls. The mean independent observer score in our study ($M = 12.18$, $SD = 6.06$), based on 17-minute samples of behavior, seems reasonably comparable to the mean Total Problems score of 9.1 ($SD = 4.1$), obtained from trained observers for a group of children that were referred for outpatient mental health or special school services, as reported by Achenbach and Rescorla.

Informant Differences

Means, standard deviations and cross-informant correlations for the videotaped child behavior ratings (DOF) are presented in Table 2. A repeated-measures ANOVA was conducted to test for differences between informants. Planned comparisons indicated that mothers reported significantly fewer behavior problems than did teachers or group-care workers (contrast: mothers vs. teachers and group-care workers: $F(1, 33) = 6.827$, $p = .013$). This result is contrary to our prediction (first hypothesis) that mothers would report more behavior problems than would professionals.

Table 2.

Means and Correlations for Different Informants on Child Behavior Ratings (DOF).

				Pearson *r*[a]		
Informant	*N*	*M*	*SD*	Teacher	Group-care worker	Independent raters
Mother						
Total observations	55	25.55	15.25			
Common observations[b]	34	25.56	14.52	.45**	.26	.45**
Teacher						
Total observations	42	29.36	16.20			
Common observations	34	31.00	16.76		.43*	.28
Aggregated observations[c]	12	28.78	13.41			
Group-care worker						
Total observations	44	35.70	20.06			
Common observations	34	34.94	20.65			.34
Aggregated observations	26	37.80	18.91			
Independent raters[d]						
Total observations	55	12.18	6.06			
Common observations	34	12.50	5.58			

[a]Interrater correlations are based on the observations common to all informants (N = 34). [b]Child behavior samples observed by all informants. [c]Obtained by averaging the child behavior ratings for each teacher or group-care worker. [d]Consensus score of the two independent raters. * $p < .05$, ** $p < .01$, two-tailed.

Familiarity with the Child

In order to test for the effect of familiarity with the child, planned comparisons were conducted for mothers vs. independent observers, as well as for teachers and group-care workers vs. independent observers (repeated measures ANOVA). We hypothesized that mothers would be likely to give more favorable appraisals of their children's behavior, and consequently report fewer behavior problems, than would the independent observers. However, results showed that mothers reported significantly more behavior problems than did the independent observers (contrast: mothers vs. independent observers: $F(1, 33) = 129{,}529$, $p = .000$). Teachers and group-care workers also reported significantly more behavior problems compared to the independent observers (contrast: teachers and group-care workers vs. independent observers: $F(1, 33) = 65.731$, $p = .000$). Thus, the results suggest that the effect of familiarity can be generalized across informants, but the direction of this effect appeared to be opposite to what we predicted in the case of the mothers.

Finally, the cross-informant correlations were only modest, ranging from .18 to .45. These correlations, obtained under controlled conditions, are comparable with the correlations

between different types of informants generally found in field research (Achenbach et al., 1987).

Trait-related Perceptual Distortions

Partial correlations between the Big Five personality traits (NEO) of the mothers, teachers and group-care workers and their child behavior ratings (DOF), while controlling for the independent observers' ratings (DOF), are presented in Table 3.

Table 3.

Partial Correlations between NEO Personality Traits and Child Ratings (DOF) controlling for Independent Ratings.

Informant	*N*	Neuroticism	Extraversion	Openness	Agreeable-ness	Conscien-tiousness
Mother	55 (55)	.10	-.04	-.14	-.05	.09
Teacher	12 (42)	.65*	-.01	-.14	-.00	.23
Group-Care Worker	26 (44)	.51**	-.51**	-.37*	-.11	-.27

* $p < .05$, ** $p < .01$, one-tailed.

Contrary to our hypothesis, no significant correlations were found between the mothers' Neuroticism, Extraversion, or Openness, and their child behavior ratings. For the teachers and group-care workers, however, significant ($p = .015$ and $p = .005$, respectively) correlations were found between their Neuroticism and their child ratings. Neuroticism explained 42% of the residual variance in the child behavior ratings in the case of the teachers, and 26% in the case of the group-care workers, after partialing out the variance shared with the independent observers. Also, significant negative correlations were found between the group-care workers' Extraversion ($p = .005$, 26% residual variance explained) and Openness ($p = .033$, 14% residual variance explained), and their child ratings. This indicates that group-care workers who were less extraverted or open were likely to report higher levels of child behavior problems. In sum, personality-related perceptual distortions were found for the teachers and group-care workers, as was hypothesized, but not for the mothers. No significant relations were found between Agreeableness and Conscientiousness and the child ratings for any of the informants.

DISCUSSION

The present study is the first to examine the impact of the personality traits of both mothers and child-care professionals on ratings of child behavior problems using standardized behavior samples. Both mothers and professionals rated videotaped samples of the behavior of a child familiar to them, which provided considerably greater control over the confounding effects of situational variability than might be accomplished in field studies. The focus of the research was on the occurrence of cross-informant differences under controlled conditions, the effect of familiarity on child ratings, and the occurrence of trait-related distortions of child behavior ratings among mothers and child-care professionals.

The results showed that mothers reported fewer behavior problems than did the professionals, that the informants who were familiar with the child reported more behavior problems than did the independent observers, and that informant Neuroticism was significantly related to higher problem ratings by the professionals, but not by the mothers. In addition, group-care workers with lower levels of Extraversion or Openness were likely to report more child behavior problems, but this was not the case for the mothers or teachers. Finally, no relations were found between Agreeableness or Conscientiousness and child ratings by any of the informants.

Familiarity with the child seems to play an important role in moderating informant differences found in our study. Both the mothers and the professionals — familiar with the child — reported significantly higher levels of problem behaviors than did the independent observers. These findings did not confirm the hypothesis that mothers are positively biased when appraising their own children as opposed to an independent observer (Kendziora and O'Leary, 1998). Other studies have also questioned the general nature of such a positive appraisal bias in mothers regarding their own children (cf. Snarr et al., 2003). These conflicting findings might at least partly be due to sampling differences, since the children in the study by Kendziora and O'Leary were younger and had less severe behavior problems than the children in our study. As already suggested by Kendziora and O'Leary, it is quite possible that the direction (positive or negative) of maternal processing biases is moderated by child difficulty level as well as by the duration of the relationship with a problem child.

As an alternative to the hypothesis that mothers are either positively or negatively biased towards their own children, our findings might also be interpreted as an "acquaintance effect" (Blackman & Funder, 1998; Funder, 1999). This alternative is supported by the fact that mothers, teachers and group-care workers all reported more behavior problems than the independent observers. Ample evidence indicates that increased acquaintance — i.e. knowing a person better and thus having greater access to information about his or her behavior — affects the accuracy of personality judgment. Acquaintance or familiarity with the target enhances the detection and utilization of behavioral cues by the observer, and thus may

account for the detection and reporting of a greater number of child behavior problems by informants familiar with the child, as was found in our study. This implies that the child reports of the mothers, teachers and group-care workers, all based on a wider experience with the child's problem behaviors, might be reflecting the actual presence of behavior problems better than the reports of the independent observers, rather than reflecting perceptual distortion.

Contrary to our expectations, mothers reported fewer behavior problems than did the child-care professionals. This finding contrasts with the results obtained in many field studies showing that mothers generally report more behavior problems than professionals (cf. Kroes et al., 2003; Youngstrom et al., 2000). One possible reason for this contrasting finding is that child behaviors may vary across situations in field studies, with children showing more behavior problems at home than in the classroom or play group, whereas the behavior samples used in our study were the same across informants. Another possible explanation is that a cognitive phenomenon known as the "anchoring-and-adjustment" heuristic played a role in creating biases in the mothers. The anchoring and adjustment heuristic is said to be descriptive when different starting points "yield different estimates which are biased toward the initial values" (Garb, 1998, p. 184). For example, a mother may be more likely to describe her child as being well adjusted in the clinic if she is used to the child's relatively more disturbed behavior at home. For the teachers and group-care workers, the child behaviors on the video might not differ substantially from the behaviors generally observed in the classroom or clinic, and therefore there might be no adjustment in their child ratings.

Our examination of the distorting effects of informant personality traits on the ratings of child behavior problems provides clear evidence for the neuroticism-related distortion hypothesis for the professionals, but not for the mothers. As already suggested in the literature (e.g., Richters, 1992; Youngstrom et al., 2000), distortion related to the psychological symptoms of the informant appears to hold for different types of informants, including teachers and other child-care professionals. Our findings are consistent with the results of other studies showing child behavior ratings to be systematically related to the psychological symptoms of peer raters (Epkins, 1994), sibling raters (Epkins & Dedmon, 1999), and university student raters (Johnston & Short, 1993). Our results also suggest that perceptual distortion is not confined to neuroticism, but may involve extraversion and openness as well.

For the mothers in our study, no evidence for trait-related distortion was found. This finding is in line with the results of other studies showing that, for instance, depressive symptoms in mothers do not produce distortions in their child behavior ratings when compared to criterion ratings (e.g., Conrad & Hammen, 1989). But, as already noted, our results contrast with a host of other studies that have claimed evidence for depression-related distortion in mothers — although the influence of situational variability has yet to be ruled out in most such cases (Fergusson et al., 1993; Kroes et al., 2003; Najman et al., 2000).

What remains to be explained is why we found distortion related to Neuroticism in the case of the professionals (and to Extraversion and Openness in the case of the group-care workers), but not in the case of the mothers. One possible explanation is that emotional factors may have played a differential role in producing perceptual distortions. For instance, Weis and Lovejoy (2002) showed that mothers reported significantly more behavior problems as compared to independent observers depending on their emotional state generated by their previous interaction with the child. They also found that maternal traits affected their child ratings, but these effects appeared to be mediated by emotional state. For the mothers in our study, the rating of videotaped child behaviors under laboratory circumstances may have differed greatly from the rating of child behaviors in real-life situations. Moreover, viewing a videotape of their child in a clinic may have created some kind of emotional distance. We assume that the impact of maternal stress generated in daily life might be reduced in this situation and, consequently, the impact of maternal traits on their ratings might have been reduced accordingly. For the professionals, on the other hand, the behaviors that were recorded in the clinic may not have differed much from those regularly observed in their daily experience with the children, nor might the rating of videotapes recorded in the clinic have created a greater emotional distance than the ratings of child behaviors in everyday life. Consequently, the rating of videotapes may have triggered the influence of personality traits to a greater extent in the case of the professionals than in the case of the mothers.

The use of two independent observers who provided criterion ratings based on exactly the same behavior samples as were rated by the mothers and professionals is a strong methodological feature of this study. The independent observers were well-trained in the use of the DOF, achieved good reliability scores during training and showed considerable overall inter-rater agreement ($r = .85$). Yet, as Richters (1992) pointed out, even when independent observers are rating the same children in the same setting, they may nevertheless use different thresholds and have different motivations, abilities and observational skills. As discussed above with regard to the acquaintance effect, the independent observers in our study had less experience with problem children than the mothers and professionals, and therefore may have underestimated the actual levels of child behavior problems. This makes it difficult to interpret the results of the group-level comparisons as reflecting either bias or accuracy. But the use of potentially lower-than-actual criterion ratings may not have greatly affected our findings regarding trait-related distortion in the teachers and group-care workers, which were based on correlational analyses.

Some caveats with respect to the interpretation of the present results may be mentioned. First, the sample size in our study was small, especially the number of teachers and group-care workers involved. Second, the informants in our study only evaluated problem behaviors, whereas other studies have found different results when evaluating positive behaviors (e.g., Youngstrom et al., 1999). Third, although our study was conducted in a clinical setting, the evidence for personality-related distortion in the case of child-care

professionals was based on viewing only 17-minute behavior samples on videotape, under controlled conditions, and therefore may not be readily generalized to clinical practice. More research is certainly needed to confirm these findings, and to specify the conditions under which perceptual distortions are likely to occur, and in which informants.

The results of the present study may have important implications for both future research and clinical practice. First, our findings suggest that the child ratings provided by professionals are not immune to distortions based on their own personality. Personality traits such as neuroticism, extraversion or openness may affect professional ratings on widely used child behavior checklists or clinician-rated measures of mood. Second, our research suggests that neuroticism in mothers may not necessarily lead to biased reports of their children's behavior problems. Under controlled conditions, mothers' ratings of their children seem unaffected by their own personality traits. Our results therefore suggest that perceptual distortion is situation-specific, rather than general.

Chapter 5

The Perception of Child Behavior Problems: The Role of Acquaintanceship, Informant Personality, and Context

Gert Kroes, Jan W. Veerman, and Eric E. J. De Bruyn

The role of acquaintanceship, informant personality, and situational context on reports of child behavior by different informants was examined within the framework of a general theory of personality judgment. Mothers and group-care workers rated videotaped behavior samples of a familiar and an unfamiliar child in the clinic and daily-life behavior of the familiar child. Independent observers also rated the videotapes. In line with the acquaintanceship hypothesis, mothers were found to perceive more behavior problems than independent observers when rating familiar children but not unfamiliar children. While maternal neuroticism was associated with maternal child reports in the context of everyday life but not in the clinic, group-care worker neuroticism was associated with their child reports in both situational contexts. The clinical and methodological implications of these findings are discussed.

INTRODUCTION

Assessment of child behavior problems greatly relies on the reports of such adult informants as parents, teachers, and other professionals who often experience the child in very different situations. Only modest agreement is invariably found between the reports of different informants, and the question that arises is obviously which informant provides the most accurate information with regard to a particular child (Loeber, Green, Lahey, & Stouthamer-Loeber, 1991). A large body of research has shown both the accuracy of any individual source of information and the degree of consensus among informants to depend on various situational and informant characteristics (Achenbach, McConaughy, & Howell, 1987; Kroes, Veerman, & de Bruyn, 2000). The aforementioned characteristics include the consistency of child behaviors across situations, the amount of behavioral information available to the informant, motivation on the part of the informant, personality of the informant, and informant familiarity with the child (Youngstrom, Loeber, & Stouthamer-Loeber, 2000).

While such situational and informant characteristics have been the subject of ample research within the domains of child assessment and personality judgment, no attempts have been made — to our knowledge — to apply the relevant theoretical concepts and findings from the study of personality judgment to the study of child assessment. In the present study, it is therefore attempted to bridge this gap with application of one of the most important theories of personality judgment — namely Funder's (1995) Realistic Accuracy Model — to the study of bias and accuracy in child assessment. In addition, we will re-evaluate some apparently contradictory findings with regard to bias in the perception of child behavior problems due to informant personality characteristics in light of the role of acquaintanceship and the context in which child behavior problems are observed.

Acquaintanceship and the Perception of Child Behavior

Some of the most contradictory results with regard to the perception of child behavior problems concern the impact of acquaintanceship or familiarity with the child. Acquaintanceship has been found to affect the perception of child behaviors in very different — and apparently inconsistent — manners. In some studies, a tendency for mothers to provide more favorable appraisals of their own children's behavior as opposed to unfamiliar children's behavior has been found (Kendziora & O'Leary, 1998). In other studies, no such differential effect of familiarity has been found (Youngstrom et al., 1999) or mixed results have been reported (Snarr, Strassberg, & Slep, 2003). In the study by Snarr et al., for

example, the mothers of oppositional boys demonstrated a negative interpretive tendency while the mothers of control children demonstrated a positive interpretive tendency but only for stimuli representing their own children as opposed to unfamiliar children. On the basis of these contradictory results, Snarr et al. suggested that the biases in maternal ratings of their own children may not be universal but depend on the severity of the behavior problems and the parent-child history.

A more general explanation for the impact of familiarity on child behavior ratings is provided by the so-called "acquaintanceship effect" (Funder, 1999). Ample evidence shows increased acquaintance (i.e., knowing a person better and thus having greater access to information about his or her behavior) to clearly affect the accuracy of personality judgments. For example, Blackman and Funder (1998) showed both interjudge consensus and accuracy, defined as self-other agreement, to be much higher for observers who had known the person being judged for an average of 14 months than for those who were unfamiliar with the target and only observed the person on video for 30 minutes. Acquaintanceship or familiarity with the target thus enhances detection and interpretation of behavioral cues. Acquaintanceship may thus account for the reporting of both greater amounts of positive and negative behaviors for children who are familiar versus unfamiliar to the observer. That is, mothers may simply be more accurate — and not biased — when reporting on their own as opposed to unfamiliar children.

The acquaintanceship hypothesis provides not only an alternative explanation for the observed variation in maternal perceptions of children's behavior but can also help us reframe one of the most frequently discussed methodological issues in research on observer biases due to informant personality traits — namely, the use of independent observers to provide criterion ratings (Richters, 1992). Most of the research on this topic has been concerned with the demonstration of bias due to maternal depression. In most of the studies, depressed mothers are found to report greater child behavior problems than both non-depressed mothers and other informants serving as criterion raters (e.g., teachers, group-care workers). As Richters has pointed out, however, none of these studies provides convincing evidence of depression-related distortion because the mothers and criterion raters invariably rate different child behaviors in different contexts. According to Richters, carefully validated and independent ratings of the same child behaviors in identical settings are needed to provide conclusive evidence of bias in the perception of child behavior. In order to meet this requirement, laboratory studies have indeed been undertaken with video ratings by independent observers not previously acquainted with the child serving as a criterion. In light of the acquaintanceship effect, however, the question is whether such independent criterion ratings actually supply an accurate report of child behavior.

To date, the few laboratory studies concerned with bias in the perception of child behavior as a consequence of informant personality traits (Kroes et al, 2005; Weis & Lovejoy, 2002; Youngstrom et al., 1999) have produced contradictory results. In the study by Weis and

Lovejoy, mothers and independent observers were asked to rate the occurrence of positive and negative child behaviors on the basis of 15-minute videotaped recordings. The mothers completed a questionnaire addressing the general occurrence of positive and negative emotions and thus maternal traits one week prior to the watching of the videotapes. Maternal mood was measured using the same questionnaire completed immediately prior to the watching and rating of the videotapes. The results showed maternal ratings of their own children to be influenced by maternal mood but not maternal personality traits. In those cases where a significant association was found between maternal personality traits and maternal ratings of their own child's behavior, the association was mediated by maternal mood. When Youngstrom et al. studied maternal reports of behavior and emotion for one's own child versus a control child and independent observer ratings using videotaped behavior samples, small but significantly positive correlations were found between maternal dysphoria and maternal reports for both their own and the control children after the ratings provided by the independent observer were taken into consideration (i.e., controlled for). Maternal dysphoria was measured using a combination of self-report rating scales — for depression, state anxiety, and trait anxiety — and actual observations of maternal emotions. When Kroes et al. recently studied the impact of the Big Five personality traits of mothers, teachers, and group-care-workers on reports of the behavior of a familiar child using 17-minute videotapes, informant neuroticism positively related to reports of child behavior problems for the teachers and group-care workers but not for the mothers. In addition, those group-care workers who were more extraverted or open reported fewer child behavior problems than those group-care workers with average levels of extraversion and openness.

In sum, the results of laboratory studies using independent criterion raters appear to mirror the mixed results obtained in field studies. Rather unintentionally, however, the results of the laboratory studies raise some doubts about the use of independent observers (i.e., individuals unfamiliar with the target) as criterion raters in the evaluation of child behavior. In all of the laboratory studies, considerable differences were found between the ratings provided by informants familiar with the children and independent observers. In the study by Kroes et al. (2005), in fact, the mothers, teachers, and group-care workers reported twice as many problem behaviors as independent judges of the same children on videotape. The results presented by Youngstrom et al. (1999) also reveal considerable differences in the ratings provided by mothers versus independent observers for negative child behaviors. Finally, in the study by Weis and Lovejoy (2002), the mothers reported more than twice as many — both positive and negative — child behaviors as the extensively trained independent observers, which led the authors to conclude that (p. 223) "mothers and observers used the scales differently, with mothers reporting higher levels of all behaviors." The differences between the ratings provided by mothers versus independent observers will be re-evaluated in light of the acquaintanceship hypothesis in the present study.

Context and the Perception of Child Behavior

In addition to acquaintanceship, the general context in which child behaviors are observed or experienced may also affect the perception of such. As Funder (1999, p. 118) has pointed out, "the process of achieving accuracy, encompasses the cognitive mechanisms of the judge, the actual attributes of the target, and the way information about the latter enters the former during transactions in the social environment." The information available with regard to a target can differ both quantitatively and qualitatively depending on the situational context. According to Funder: (1999, p. 128), contexts "differ not just in how much expression they allow but in which attributes of personality they provide an opportunity to express." A child may display certain behaviors at home but not in other contexts, for example. Moreover, specific characteristics of the individual interacting with the child may elicit specific behaviors on the part of the child, and the behavior displayed by a particular child may thus depend on the interactants in a particular context. Therefore, a child interacting with his or her mother at home may display different behavior than the same child interacting with his or her friend at home, or elsewhere.

A given context may also elicit a particular response set or particular emotions on the part of an informant and thereby influence his or her judgments. As Weis and Lovejoy (2002) have demonstrated, negative emotions evoked in the judgment context can clearly lead to biased reports of child behavior. Within the scope of this study, both types of contextual influence are important, but we will focus primarily on the role of situational context in constraining the type and range of behaviors available to the judge.

Several mechanisms have been described by which depression in mothers may lead to increased behavioral and emotional problems in their children (Goodman & Gotlib, 1999). As Goodman and Gotlib note, some of these increased behavioral problems may only occur during the interactions between the depressed mothers and their children. As a consequence, depressed mothers are likely to actually perceive more context-specific negative child behaviors than other observers judging the child in a different context. But it is also possible that depression in mothers leads to a biased report of their children's negative behaviors. The question of bias or accuracy in the reports of depressed mothers can thus be reframed as follows: Do children of depressed mothers actually express more behavior problems during interactions with their mothers or do depressed mothers report more child behavior problems than actually exist?

In research on biases related to informant personality characteristics (e.g., anxiety, depression), the context in which the behavior is sampled (i.e., in actual interaction with the informant or not) may thus play a crucial role. What holds for anxiety and depression is also likely to hold for such related personality characteristics as dysphoria and neuroticism, with the latter constituting facets of anxiety and depression within the Big Five model of personality (Costa & McCrae, 1992). In other words, some of the contradictory findings in the

aforementioned laboratory studies may be due to the intricate interaction between informant personality characteristics and sampling context. Along these lines, Youngstrom et al. (1999) found a significant relation between maternal dysphoria and maternal ratings of video recordings of their children performing a frustrating task in the presence of the mother while Kroes et al. (2005) found no relation between maternal neuroticism and maternal ratings of video recordings of their children playing in the clinic with their mothers not present.

The present study

In light of the contradictory nature of the research findings concerned with informant judgments of child behavior, both the influence of acquaintanceship and context on the perception of child behavior were examined in the present study. In order to provide convincing evidence of an acquaintanceship effect, ratings of familiar and unfamiliar children provided by mothers, group-care workers, and independent observers using the same behavioral rating scale were compared. On the basis of the relevant acquaintanceship literature (Blackman & Funder, 1998; Funder, 1999), we expected mothers and group-care workers to report higher levels of problem behaviors for *familiar* children than independent observers but similar levels of problem behaviors for *unfamiliar* children.

With respect to the role of judgment context, the association between informant personality traits and judgments of child behavior were examined using the same informants in a laboratory setting and everyday life. The judgments of child behavior in the laboratory setting were based on videotaped behavior samples recorded in the clinic without any interaction between the informant and the child; the judgments of child behavior in everyday life were based on child behavior observations during daily interactions between the informant and the child. We hypothesized that informant neuroticism would influence the child problems expressed during interaction with the informant. If the informant accurately perceives and reports the child behavior evoked by his or her neuroticism, however, higher levels of informant neuroticism can be expected to be associated with higher ratings of problem behaviors in everyday life (i.e., in interaction with the informant) but not in laboratory settings (i.e., in the absence of any interaction with the informant). If informant personality generally distorts judgments of child behavior but does not further influence actual child behavior, higher levels of informant neuroticism can be expected to be associated with higher ratings of problem behaviors in *both* the everyday life and laboratory settings. Given that the influence of informant extraversion on child behavior in everyday life is as yet unknown, a similar set of — largely exploratory — analyses was undertaken to examine the influence of informant extraversion on child behavior and judgments of child behavior in both everyday life and a laboratory setting.

METHOD

Participants

The subjects were 55 elementary school-aged children (43 boys and 12 girls) receiving treatment at de Waarden, a Dutch clinic for the treatment of youth with emotional, behavioral, and learning problems. During two consecutive years, all of the children attending the elementary school of the clinic were recruited along with their mothers to participate in the study. From an initial pool of 63 children whose mothers had agreed to take part in the study, eight cases were omitted due to incomplete data (3), premature termination of treatment (2), maternal illness (2), or withdrawal of consent to participate by the mother (1). The mean age of the children in the remaining sample of 55 children was 8.9 years (*SD* = 1.9, range of 6 to 13 years). All of the children were diagnosed with psychiatric disorders, such as Attention Deficit Hyperactivity Disorder, Pervasive Developmental Disorder, Oppositional Defiant Disorder, or Conduct Disorder according to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (American Psychiatric Association, 1994). None of the children met the criteria for mental retardation. The mean age of the mothers was 38.5 years (*SD* = 4.8, range of 25 to 52 years). The occupational status of the mothers was categorized using the Social Demographic Inventory (SDI; van Westerlaak, Kropman, & Collaris, 1975). Rated along a 6-point scale ranging from (1) unskilled labor to (6) academic career, the median SDI score was 3. The mothers signed consent forms and received a gift with a value of 15 euros after participation.

The 55 children participating in the study were involved in either residential (47%) or day treatment (53%). Their group-care workers were also asked to participate in the study. And in order to establish some consistency in the familiarity of the caregivers with the children, those who did not know the target child for at least six months were excluded. Other group-care workers were not included due to job change, illness, or refusal to participate. The final sample included 26 (80%) of the group-care workers asked to participate. This resulted in 44 separate group-care worker-child dyads as the caregivers sometimes worked with more than one of the 55 children participating in the study. The group-care workers thus rated the behavior of 1.69 children on average (*SD* = 1.05, range of 1 to 5). The mean age of the group-care workers was 33.0 years (*SD* = 6.77), and 81% was female. All of the group-care workers had a college degree in education. The professional experience of the group-care workers ranged from 1 to 20 years (*M* = 6.8, *SD* = 5.1). The group-care workers also received a gift with a value of 15 euros for their participation.

Measures

NEO Five Factor Inventory (NEO-FFI)

Informant personality traits were assessed using the NEO-FFI (Costa & McCrae, 1992; Dutch version by Hoekstra et al., 1996). The NEO-FFI is the short form of the Dutch NEO-PI-R (Hoekstra et al., 1996), which is an authorized translation of the Revised NEO Personality Inventory by Costa and McCrae (1992). The NEO is a widely used and well-validated self-report measure of the Big Five personality traits of Neuroticism, Extraversion, Openness, Agreeableness, and Conscientiousness. The NEO-FFI consists of 60 items to be rated along a 5-point scale. The short form takes 10-20 minutes to complete. The norms for the Dutch NEO-FFI were obtained from a general population sample (N =2415). The NEO-FFI raw factor scale scores are transformed into stanine scale scores (with a mean of 5 and a standard deviation of 2) using Dutch population norms.

Child Behavior Checklist (CBCL)

The Dutch translation of the CBCL (Achenbach, 1991; Verhulst, van der Ende, & Koot, 1996) was used to assess child behavior problems in everyday life. Mothers and group-care workers rated the presence of 118 behavioral, emotional, and social problems for the preceding six months. Each item was rated as 0 *(= not true*), 1 (= *somewhat or sometimes true*), or 2 *(= very or often true*). The sum of all the items constitutes the Total Problems score, which we used in our analyses. Confirmatory factor analysis of the CBCL results for the group-care workers showed the CBCL factor model found for the parental judgments to also fit the judgments of the group-care workers (Albrecht, Veerman, Damen, & Kroes, 2001). Only the raw scores were used in the analyses for both the mothers and the group-care workers.

Direct Observation Form (DOF)

The DOF (Achenbach, 1986; Dutch version by Ferdinand & Verhulst, 1998) was designed to assess the behavior problems observed in — among other settings — classrooms and group activities. The DOF is easily used by teacher aids and research assistants after training by an experienced observer (Achenbach & Rescorla, 2001). The DOF consists of 96 items, 72 of which have counterparts in the CBCL. Each item is rated along a scale ranging form 0 (= no observed occurrence of the behavior) to 3 (= definite occurrence with severe intensity or a duration of three or more minutes). The sum of all the items constitutes the Total Problems score, which we used in our analyses.

The DOF is normally completed after 10 minutes of live observation. For the present study, which included the rating of videotapes, an alternative assessment procedure was

followed. Rather than complete the DOF form, the informants were asked to sort a deck of 96 cards with the DOF items listed separately on them in two steps directly following the viewing of the videotape. In the first step, the mothers, group-care workers, and independent observers were asked to select those DOF items that they definitely did not observe on the videotape. The cards containing these items were then removed from the deck and assigned a score of 0 by the interviewers. In the second step, the informants were asked to sort the remainder of the cards into three piles reflecting the DOF rating scores of 1, 2, or 3. Written descriptions of the rating scores were provided to facilitate the sorting process. The interviewers subsequently transferred the results of the sorting procedure onto the DOF form. This two-step assessment procedure was designed to attain scoring consistency across the different informants by asking them to make very deliberate decisions about each item and providing written descriptions to anchor their decision making.

Procedure

Videotapes of the children were made in their play groups at the treatment center using a script to both standardize the behavior samples and elicit a range of naturalistic child behaviors. The video script was tested in a previous study (Muntinga, 1999). The play group was supervised by an assistant group-care worker who was instructed to intervene only when deemed necessary to maintain order in the group. A 17-minute videotaped behavior sample was obtained for each of the target children. To attain a behavior sample of an unfamiliar child, two additional 17-minute recordings were made in the same setting of a boy and a girl not known to the mothers or the group-care workers participating in the study.

The mothers and group-care workers were asked to watch the videotapes of a familiar child and an unfamiliar child individually and then assess the behavior problems observed on the videotapes. The order of presentation for the videotapes of the familiar and unfamiliar children was counterbalanced, and the sex of the unfamiliar child was matched to the sex of the familiar child. The sessions were all conducted at the clinic by trained research assistants (N = 5) who were unfamiliar with the target children. The research assistant conducting the session with a particular informant was also unaware of the ratings provided by the other informants for the same child. During the sessions, the mothers and group-care workers first completed the NEO-FFI. Thereafter, they viewed the videotapes and were asked — immediately following each tape — to assess the behavior problems observed on that tape using the DOF.

The CBCL was completed at home by the mothers and returned by mail prior to the session at the clinic. The group-care workers had previously completed the CBCL as part of their regular treatment activities.

The independent criterion measures were made by two observers who individually viewed the videotapes and completed the DOF for the 55 target children and 2 control children. The independent observers were undergraduate psychology students with previous training in child observation and assessment. The independent observers were given additional training on the DOF by an experienced observer until the inter-rater Kappas across the DOF items reached at least .80 for a set of five pilot videotapes. The final inter-rater agreement between the independent observers for the videotapes of the target children (N = 55) was .85. After their initial — individual — rating of the videotapes, the independent observers discussed any discrepancies until consensus could be reached. The independent observers were instructed to review the videotapes as many times as they felt necessary, and the agreed upon scores were used in all of the statistical analyses. Both the mothers, group-care workers, and independent observers were instructed to rate only those behaviors that they had actually observed on the videotape.

Data Analysis

The effects of acquaintanceship were evaluated in one-way repeated-measures ANOVAs with informants (i.e., mothers, group-care workers, independent observers) as the independent variable and DOF ratings of child behavior problems as the dependent variable. Planned contrasts were tested between mothers versus independent observers and group-care workers versus independent observers. Separate analyses were conducted for the ratings of the familiar versus unfamiliar children.

The association between informant neuroticism and extraversion, on the one hand, and child behavior ratings in everyday life and the laboratory setting, on the other hand, was analyzed in multiple regression analyses. Two different models were tested. First, the influence of informant personality traits on child behaviors in everyday life was tested by examining the contribution of NEO personality traits to the explanation of the variance in the CBCL ratings provided by the mothers and group-care workers (model I). Separate regression models were calculated for the mothers versus group-care workers and for neuroticism versus extraversion. Second, the contribution of NEO personality traits to the explanation of the variance in the laboratory DOF ratings provided by the mothers and group-care workers was examined (model II).

RESULTS

Descriptive Data

The means and standards deviations for the child behavior ratings provided by the mothers and group-care workers in everyday life (i.e., the CBCL ratings) and the mothers, group-care workers, and independent observers in the laboratory setting (i.e., the DOF ratings) are presented in Table 1.

Table 1.

Raw scores and ANOVA Results for Child Behavior Ratings by Different Informants in Different Contexts.

				Repeated measures (GLM)			
Informant	*N*	*M*	*SD*	*Planned contrast*	*df*	*F*	*p*
Familiar child, daily life ratings (CBCL)							
Mothers	55	54.98	21.89	Mothers vs. group-care workers	1,54	26.13	.000
Group-care workers	55	38.35	20.62				
Familiar child, laboratory ratings (DOF)							
Mothers	55	25.55	15.25	Mothers vs. independent observers	1,43	47.62	.000
Group-care workers	44	35.70	20.06	Group-care workers vs. independent observers	1,43	68.95	.000
Independent observers	55	12.18	6.06				
Unfamiliar child, laboratory ratings (DOF)							
Mother	55	19.09	14.03	Mothers vs. independent observers	1,25	0.02	.902
Group-care workers	26	34.62	20.34	Group-care workers vs. independent observers	1,25	19.01	.000
Independent observers	55[a]	17.65	3.0				

[a] The sexes of the unfamiliar children were matched to the sexes of the familiar children rated by the mothers.

As already noted, only the raw scores have been used in the analyses. The mean CBCL Total Problems score for the mothers was well above the Dutch population mean of 21.3 for boys four- to eleven-years old and 19.2 for girls four- to eleven-years old (Verhulst et

al. 1996). The mean CBCL score for our sample corresponds to a T-score of 69, which indicates severe behavior problems. The mean CBCL scores for the group-care workers were also well above the Dutch population mean but should be considered with caution as the Dutch parental norms — in the absence of normative data for group-care workers — were used for comparison.

The cross-informant Pearson correlation between the CBCL ratings provided by the mothers and group-care workers (.36) was significant and in line with the correlations that are generally found to occur between different informants (e.g., an average correlation of .24 between the judgments of parents and mental health workers; Achenbach et al., 1987). As can be seen from Table 1, the mean level of behavior problems observed by the mothers in daily life differs significantly from the mean level of behavior problems observed by the group-care workers during their daily experiences with the children. This finding thus confirms the general observation that mothers tend to report more child behavior problems than professionals (Youngstrom et al., 2000).

The DOF behavior problem scores obtained in the laboratory setting were compared to the DOF scores presented in the *Manual for the ASEBA school-age forms & profiles* (Achenbach & Rescorla, 2001). In the Manual, an average Total Problems score of 9.1 (SD = 4.1) is reported for trained observers rating 10-minute samples of the classroom behavior of referred children. In our study, the independent observers assigned an average Total Problems score of 12.18 to the videotaped behavior samples for the children familiar to the mothers and group caregivers and an average Total Problems score of 17.65 to the videotaped behavior samples for the unfamiliar children.

The mothers in our study reported average levels of neuroticism on the NEO-FFI (N = 55; M = 5.1, SD = 1.9; stanine scale scores) when compared to the general Dutch population (M = 5, SD = 2), and also average levels of extraversion (M = 5.2, SD = 2.0). For the group-care workers, the mean NEO-FFI neuroticism score of 4.08 (N = 26; SD = 1.5) was slightly lower when compared to the general Dutch population and the mean NEO-FFI extraversion score of 6.35 (SD = 1.7) was slightly higher.

Acquaintanceship

The results of the repeated measures ANOVA to evaluate the role of acquaintanceship in the laboratory ratings of child behavior problems are also presented in Table 1. As can be seen, the laboratory ratings provided by the mothers significantly differ from the ratings provided by the independent observers for the familiar children but not the unfamiliar children. In the case of the group-care workers, however, the laboratory ratings significantly differ from the ratings provided by the independent observers for *both* the familiar and

unfamiliar children. The acquaintanceship hypothesis is thus confirmed by the mothers in our study but not the group-care workers.

Informant Personality and Context

The results of the hierarchical multiple regression analyses to examine the influence of informant personality traits on child behaviors and judgments of child behaviors in different contexts are presented in Table 2.

Table 2.

Results of Hierarchical Regression Analyses with Informant Personality Traits predicting Familiar Child Behavior Ratings in Different Contexts.

Predictor	*N*	*R*	R^2	β	*p*
Model I. Daily life ratings (CBCL) as dependent variable, single predictor					
Mothers	55				
Neuroticism		.42	.18	.42	.00
Extraversion		.15	.02	-.15	.29
Group-care workers	44				
Neuroticism		.52	.27	.52	.00
Extraversion		.07	.01	.07	.65
Model II. Laboratory ratings (DOF) as dependent variable, single predictor					
Mothers	55				
Neuroticism		.07	.00	.07	.63
Extraversion		.05	.00	-.05	.74
Group-care workers	44				
Neuroticism		.47	.22	.47	.00
Extraversion		.32	.10	-.32	.03

Significant correlations were found between the informant personality trait of neuroticism and child behavior ratings in everyday life (model I). Higher levels of neuroticism were found to be associated with higher levels of reported child behavior problems for both the mothers and group-care workers as informants. In the laboratory setting (model II), significant associations between both the personality characteristics of neuroticism and extraversion, on the one hand, and the ratings of child behavior problems, on the other hand, were found but only for the group-care workers and not the mothers. The hypothesis

that informants accurately report behavior problems associated with their neuroticism was thus confirmed for the mothers but not for the group-care workers. More generally, our results show the impact of informant personality on ratings of child behavior to vary depending on the context in which the judgments are made.

DISCUSSION

The focus of the present study was on the effects of acquaintanceship, informant personality, and context on the perception of child behavior. Drawing on Funder's (1999) theory of personality judgment, we created a research design that allowed us to examine whether acquaintanceship affects the perception of child behavior problems or not. In addition, it was possible to investigate the influence of informant personality traits and situational context.

Mothers were indeed found to perceive more behavior problems than independent observers when rating their own children's behavior but not the behavior of unfamiliar children. Mothers perceived the same levels of problem behaviors for unfamiliar children as independent observers of the same children. These results provide strong evidence for an effect of acquaintanceship. However, the group-care workers were found to perceive significantly more behavior problems for both the familiar and unfamiliar children when compared to independent observers, which suggests that mechanisms other than or in addition to acquaintanceship must have influenced the perception of child behavior problems by the professionals.

The situational context in which the child was observed appeared to play a critical role in the associations between informant personality and ratings of child behavior problems. For the mothers in our study, neuroticism related to their ratings of child behavior problems in everyday life but not in the laboratory. For the group-care workers, neuroticism was found to be associated with their perceptions of child behavior problems in both contexts.

Our finding that mothers report significantly more behavior problems relative to independent observers for familiar children in the laboratory but not unfamiliar children was predicted by the acquaintanceship hypothesis. As Blackman and Funder (1998) have shown, an acquaintance appears to be a more accurate judge of an individual's personality than a stranger. Within the context of the present study, this means that mothers may be more accurate judges of the problem behaviors of their children than independent observers unfamiliar with the children. The conclusion that mothers are more accurate and thus not biased reporters of their own children's behavior problems — at least under controlled conditions — is further supported by the fact that we did not find significant relations between maternal personality traits and their laboratory ratings of child behavior problems.

Rather surprisingly, the group-care workers — who were acquainted with the children in our study — reported approximately the same levels of problem behavior for familiar versus unfamiliar children. For both familiar and unfamiliar children, group-care worker laboratory ratings appeared to be higher than the independent observer ratings of the same children. There are several possible explanations for this finding. First, the group-care workers presumably have more experience with the evaluation of problem behaviors in clinically referred children than independent observers and are therefore better prepared to detect and interpret minor behavioral cues (cf. Funder, 1999, for the role of experience in judging people). Group-care workers may thus be more accurate than independent observers (and mothers) when evaluating the behavior of referred children. However, as Garb (1998) has noted, clinicians are also known to be influenced by the context during the evaluation of psychiatric problems. Once an individual is admitted to a clinic, thus, he or she may be judged differently, which means that the reports of the group-care workers in our study may have been biased — and not more accurate — for both the familiar and unfamiliar children admitted to the clinic.

Another possible explanation for the finding that group-care workers reported more behavior problems for the unfamiliar children than both the independent observers and the mothers in our study may be a predisposition to provide socially desirable responses in the laboratory situation in particular). Rosnow and Rosenthal (1997) have described, in a series of experiments, how participants comply with the perceived demands of an experimental task in order to be a socially desirable or good subject. For the inferential and interpretative activities of the participant in an experiment to be biased, a number of response cues "must be perceived, and the participant must be motivated and able to respond to them" (Rosnow & Rosenthal, 1997, p. 78). For the group-care workers in our study, the three factors that may lead to response bias according to Rosnow and Rosenthal may have been present in the laboratory situation in particular. The group-care workers may have perceived the rating of the behaviors of referred children as an invitation (i.e., cue) to demonstrate their professional skills. The group-care workers were presumably motivated to do this and perform well. And the group-care workers were also in a position to respond to these demands by reporting numerous problem behaviors. Taken together, these factors may have elicited a response bias on the part of the group-care workers but not the mothers or the independent observers in the same laboratory setting. This explanation is further supported by the fact that the group-care workers reported significantly less behavior problems than the mothers in everyday life (i.e., in the absence of the aforementioned cues and circumstances).

In our study, the situational context appeared to play a role in the perception of child behavior by different informants. Our finding of an association between maternal neuroticism and maternal reports of child behavior in everyday life but not in the laboratory suggests actual differences in child behavior at home in interaction with the mother and child behavior in the clinic. In line with the findings of Goodman and Gotlib (1999), moreover, it is certainly

likely that children behave differently at home, under the influence of maternal neuroticism, than in other situations. However, it cannot be ruled out that the association between maternal neuroticism and report of child behavior at home does not reflect at least some informant bias. In light of the research by Weis and Lovejoy (2002), it can be argued that maternal mood may have mediated the effects of maternal neuroticism on the interpretation of child behaviors at home but not in the laboratory. The maternal emotions generated by mother-child interactions in the home may be very different than the emotions experienced in the laboratory. To obtain conclusive evidence with regard to the distortion of perceptions of child behavior within the home in particular, independent ratings by observers familiar with the child behaving in the same context as rated by the mothers (i.e., in the home) are needed.

We can only speculate about why neuroticism related to the ratings of child behavior problems in both everyday life and the laboratory setting for the group-care workers but not the mothers. One possible explanation is that the everyday life and laboratory contexts were more alike for the group-care workers than for the mothers. The videotaped behavior samples were recorded in the room where the group-care workers meet on a daily basis with the children, which meant that the videotaped behaviors used in the laboratory may have resembled the behaviors observed in daily life and thus elicited the same response set on the part of the group-care workers as in the daily life context. Moreover, it is possible that — in addition to a predisposition on the part of the group-care workers to provide socially desirable responses within the laboratory context — neuroticism *increased* their predisposition to comply with the perceived demands of the laboratory context and thus report a greater number of problem behaviors.

Our findings have some important clinical and methodological implications. The conclusion that mothers may be more accurate judges of at least their own child's behavior problems and that maternal neuroticism is not associated with maternal reports of child behavior in at least the laboratory setting proves that Richters (1992) was right in stating that conclusive evidence of trait-related biases in maternal perceptions has yet to be delivered. However, the conclusion that mothers may be more accurate judges than independent observers also seriously questions the use of criterion judgments as provided by independent observers for the study of trait-related biases, as proposed by Richters. In light of the results of our study and the other laboratory-based studies reviewed above, we therefore suggest that Richters' methodological conditions should be supplemented with the requirement that criterion ratings be provided by informants familiar with the children being observed.

Finally, some possible caveats with regard to the results of our study should be pointed out. First, the sample size was relatively small, particularly for the group-care workers. Second, the informants in our study evaluated only problem behaviors. Nevertheless, the present results underscore the complex and intricate nature of child behavior assessment, just as personality assessment (cf. Funder, 1995). Paraphrasing Funder, we may thus conclude that child assessment has become more complex but also more realistic.

Chapter 6

Summary and Discussion

FINDINGS

The focus of this dissertation was on the question of who provides the most accurate information with regard to the behavior of a child under which circumstances (see Chapter 1). In clinical practice, information on the functioning of a particular child is commonly provided by such different types of informants as the parents, teachers, other professionals, and the children themselves. The agreement between the child behavior judgments provided by different types of informants has been reported to be quite low, however. One explanation for the low agreement between the different types of informants may lie in the fact that they are often exposed to behavior occurring under very different circumstances. That is, child behavior problems may be situation specific. A second explanation for the low agreement between the different types of informants may lie in the characteristics of the informants themselves. That is, the information provided on child behavior problems may also be informant specific.

In Chapter 2, the state of affairs with regard to research on the perception of child problem behavior was reviewed. There is considerable debate about the influence of parental psychopathology and particularly maternal depression on judgments of child behavior. And it is very possible that other personality characteristics may also bias parental judgments of child behavior although it is generally accepted that the evidence for such a bias is weak. The lack of criteria to evaluate the validity of child behavior judgments is a major problem. Given that each informant has his or her own perspective on the behavior of a child, the best way to proceed is to gather as much information as possible with regard to the child's behavior. And while several procedures have been proposed to structure the information gathered from different sources, none of the methods has proved superior as yet.

In order to improve the assessment of child problem behavior, we suggested that all of the factors that appear to influence perceptions and judgments of child behavior be subsumed within a single, more general, model of social perception. The Realistic Accuracy Model (RAM; Funder, 1995) includes both situational and informant factors and incorporates information from the domains of social psychology and personality judgment. We argued that

this model provides a relevant and useful theoretical framework for the study of child assessment as well.

In Chapter 3, the first of three empirical studies designed to investigate the accuracy and distortion of perceptions of child problem behavior was presented. The associations between various types of maternal psychopathology and maternal reports of internalizing and externalizing child behavior problems for a clinical sample of 68 boys between the ages of 6 and 12 years were examined using the Symptom Check List (SCL-90), Child Behavior Check List (CBCL), and Teacher's Report Form (TRF). Given that an absolute standard for the identification of child behavior problems does not exist, the reports of both teachers and group-care workers were taken to provide the most accurate information and thus as a criterion measure for the evaluation of the maternal reports. Multiple regression analyses showed substantial correlations between various types of maternal psychopathology and report of internalizing child behavior problems after control for that variance explained by the criterion ratings. In other words, mothers with higher levels of psychopathology were more likely to report internalizing problem behaviors for their children than mothers with lower levels of psychopathology even after the levels of internalizing problem behavior reported by the teachers and group-care workers were taken into consideration. Only small or insignificant correlations were found for the report of externalizing behavior problems. The distorted reporting of primarily internalizing child behavior problems is consistent with social attribution theory, which predicts greater distortion for the observation of more ambiguous stimuli.

In the first study, criterion ratings provided by multiple informants were used to be compared with the maternal ratings of child problem behavior. However, the different types of informants observed and rated the children's behavior under very different circumstances (i.e., at home, in the classroom, within the living group). The variation in the child behavior ratings provided by the different types of informants could therefore be due to informant characteristics such as maternal psychological symptoms, but it could just as easily be ascribed to situation characteristics. That is, the children of depressed or anxious mothers may actually *display* greater behavior problems at home than at school or in the clinic. To control for variation in child behavior ratings due to situational factors, standardized behavior samples are needed for observation by all informants.

In Chapter 4, a second study using videotaped samples of child behavior in the clinic for 55 children between the ages of 6 and 12 years to provide standardized samples for observation by different informants was presented. The child behavior ratings provided by the mothers, teachers, and group-care workers were compared to each other. In addition, the effects of the personality traits of the mothers, teachers, and group-care workers — assessed using the NEO Five Factor Personality Inventory — on their ratings of child problem behavior were examined. All of the informants thus watched and rated the same 17-minute

videotaped behavior sample for a familiar target child. In addition, independent observers unfamiliar with the children provided criterion ratings of child behavior problems.

The results showed the mothers to report fewer behavior problems than the teachers or group-care workers; those informants familiar with the child to report more behavior problems than those informants not familiar with the child (i.e., the independent observers); and higher levels of informant neuroticism to be related to higher ratings of child behavior problems in the case of the teachers and group-care workers but not in the case of the mothers. In addition, those group-care workers who were more extraverted and open were less likely to report child behavior problems than those group-care workers with normal levels of extraversion and openness. Finally, no relations were found between agreeableness or conscientiousness and ratings of child behavior. The results of the second study thus suggest that professionals working with children are not immune to biased perception.

The effects of familiarity (or so-called acquaintanceship) on reports of child behavior were further explored in our third and final study, which is described in Chapter 5. Funder's (1995) theory of personality judgment was adopted to analyze the role of acquaintanceship and reconsider some methodological problems associated with previous research on informant biases. In addition, the role of differing situational contexts — i.e., the home for mothers, the living group for group-care workers, and videotaped behavior samples recorded in the living group for both types of informants — in the perception of child behavior and possible distortions of such was examined. The mothers and group-care workers rated videotaped samples of the behavior of both a familiar and an unfamiliar child, on the one hand, and the daily-life behavior of the familiar children either at home (for the mothers) or in the clinic (for the group-care workers), on the other hand. Independent observers also rated the videotapes.

In line with the acquaintanceship hypothesis (Funder,1999), mothers perceived greater behavior problems than independent observers but only when they rated familiar as opposed to unfamiliar children. Group-care workers, in contrast, reported greater behavior problems than independent observers when rating both familiar and unfamiliar children. In addition, maternal neuroticism was clearly associated with maternal reports of child problem behavior but only within the context of everyday life and not within the context of the clinic (i.e., not for the videotaped samples of their own children in the clinic) while group-care worker neuroticism was clearly associated with their reports of child problem behavior in both contexts (i.e., when judging their everyday interaction with the children and the videotaped behavior samples for the same children).

CONCLUSIONS

The most serious methodological problem encountered in research on the accuracy of judgments of child behavior problems is the lack of a golden standard. To determine the

accuracy of a judgment, an independently validated criterion measure is needed. However, as things now stand, the accuracy of the judgment provided by a particular informant "inevitably must be evaluated against judgments by informants who themselves may be inaccurate for unknown reasons and to an unknown extent" (Richters,1992, p. 490).

In the laboratory studies presented and reviewed in this dissertation, trained observers unfamiliar with the children in question were initially used to provide independent criterion ratings. The use of trained observers unaware of both the child's history of problem behavior and maternal psychological symptoms may indeed be an important methodological requirement in the study of perceptual biases but, even within the laboratory context, the validity of the judgments of even such observers may not be 100%. The task of the researcher, then, is to demonstrate greater validity for the ratings provided by the independent observers relative to the ratings provided by mothers and other informants. As ample research has demonstrated an effect of acquaintanceship on social perceptions (Funder, 1999), however, it is very possible that just the opposite may actually hold in many situations. That is, ratings of behavior provided by informants *familiar* with the subjects in question may be superior to ratings provided by independent observers unfamiliar with the subjects. That is, independent criterion ratings may actually be less valid than the ratings provided by mothers, teachers, or group-care workers.

The results of our laboratory studies showed the mothers and professionals familiar with the children in the studies to report twice as many problem behaviors as independent observers unfamiliar with the children. If — in keeping with the acquaintanceship hypothesis — acquaintances are indeed assumed to be more accurate judges of behavior than strangers (Blackman & Funder, 1998), the present findings seriously question the use of independent criterion raters for behavioral observation and evaluation. When the results of the present studies are considered in conjunction with the results of other laboratory-based studies (cf. Weis & Lovejoy, 2002), moreover, it can be suggested that Richters' methodological conditions for research on informant biases should be supplemented with the requirement that criterion ratings be provided by informants *familiar* with the children being observed.

The present results also *suggest* that — at least under controlled circumstances — the psychological characteristics of mothers may also not lead to biased perceptions of their own children's behavior. Within the context of everyday life, however, significant relations were found between maternal symptoms of psychopathology and their ratings of child problem behavior. Despite efforts to provide a conclusive answer to the question of whether this association between maternal symptoms and maternal reports of child behavior in everyday life reflects a biased or accurate perception of child behavior, the question remains unanswered. To obtain more conclusive evidence with regard to the accuracy or distortion of perceptions of child behavior within the home in particular, two methodological requirements should certainly be met. First, the mothers and criterion raters should observe and evaluate the same child behaviors under exactly the same circumstances (i.e., within the home). To meet

this requirement, videotapes of the child's behavior within the home might be used rather than behavioral questionnaires or home observations. Second, independent observers familiar with the child being evaluated should be used. The maternal ratings of the same child behaviors can then be compared to the ratings of observers who are also familiar with the child. Finally, any differences between the ratings provided by different informants should certainly be analyzed in relation to the personality characteristics of the informants themselves.

The integration of those situational and informant characteristics that influence the perception of child behavior within a single model, as provided by Funder's (1995) Realistic Accuracy Model (RAM), can help structure the process of assessing children's behavior. Funder's Realistic Accuracy Model also proved useful for the generation of some new hypotheses as well as for the re-evaluation of some old truths.

The results of our studies underscore the complex and intricate nature of social perception and, along the lines of Funder, we can thus conclude that research on the accuracy of child reports has become not only more complex but also more realistic. Although Funder's (1995) RAM proved to be quite useful for the design of our studies, further research might also benefit from other models within the domain of social perception such as the Social Relations Model (SRM; Kenny & LaVoie, 1984). Whereas the RAM presents a stimulating heuristic research model, the SRM provides a clear-cut set of methods for further application to the analysis of perceptions of child behavior problems within the family as a whole. A unique feature of the SRM model is, namely, that each and every subject serves as both a judge and a target, which allows the following questions to be examined: (a) To what extent does a given family member display a general perceptual tendency (i.e., tend to perceive all other family members in a particular manner)? (b) To what extent are a given family member's perceptions of another family member shared by other members of the family? And (c) to what extent are a given family member's perceptions of another family member unique? The SRM also appears to be particularly well-suited for further analysis of the relative contributions of the personality traits of various family members in addition to the contributions of general family characteristics to the perception of child behavior problems within the home (e.g., see Delsing, 2004). In his discussion of the SRM approach, Funder (1999) indicates appreciation for the strengths of the approach but also points out some issues that may hamper the interpretation of SRM outcomes. In future research, therefore, it needs to be explored just how the RAM and SRM can be integrated into a more general research paradigm to then unravel the intricacies of the perception of child behavior and child problem behavior.

Our investigation of the perception of child problem behavior started with the question of just who provides the most accurate information with regard to a child's behavior in which contexts. Based on a review of the relevant literature, we were initially inclined to think that the personality traits of mothers could bias their perceptions of their children's behavior at times. We were also inclined to view such professionals as teachers and group-care workers

as the most accurate informants with regard to children's — albeit often situation-specific — problem behavior. As things turned out, the mothers in our studies proved to be the most accurate judges of their own children's behavior, at least under controlled circumstances, and the judgments provided by the professionals appeared to be consistently confounded by their personality traits. Self-confidence or its counterpart, neuroticism, in teachers and group-care workers as well as extraversion and openness in group-care workers were found to influence their perceptions of child behavior in the laboratory. Moreover, self-confidence or its counterpart, neuroticism, in group-care workers appeared to influence their perceptions of child behavior in the living group.

It should be emphasized that the results of the present studies do not supply evidence for the complete validity of the judgments provided by any one type of informant. The question of whether more self-confident and extraverted professionals provide more accurate judgments than neurotic or introverted colleagues therefore remains to be answered. As studies of the characteristics of good judges have shown (Funder, 1999), however, we can assume self-confidence and extraversion to be among the personality traits most closely associated with accuracy of judgment.

Almost twenty years ago, Achenbach et al. (1987) concluded their famous and widely cited meta-analysis of 119 child report studies with the statement that reports of child and adolescent behavior problems can clearly vary depending on not only the informant but also the situation. Achenbach et al. (p. 228) further suggested that clinical assessment should be "geared" to these realities and proposed a multiaxial assessment system that included different informants for this purpose. The use of different informants was primarily motivated at that time by the need to include the different aspects of a child's behavior, which could clearly vary across situations. That is, the context was at issue and not the informant. Since 1987, many studies have examined the role of the informant and mothers, in particular, in the perception of child behavior, but professionals are still considered the experts and thus taken to be the providers of the most accurate information. The results of the studies reported on in this dissertation suggest that the accuracy of professional child reports should not be taken for granted. For the assessment of child behavior problems, this means that the suggestion that as much information should be gathered about the child's functioning under different circumstances and from the perspectives of as many different informants as possible still holds. It should be added, however, that the personalities of the informants must also be taken into consideration as well as the context in which the child problem behavior is observed.

References

Achenbach, T.M., & Edelbrock, C.S. (1983). *Manual for the Child Behavior Checklist and revised profile.* Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1986). *The Direct Observation Form of the Child Behavior Checklist* (Rev. ed.). Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T. M. (1991). *Integrative Guide for the 1991 CBCL/4-18, YSR, and TRF Profiles*. Burlington: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles.* Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Achenbach, T. M., McConaughy, S. H. & Howell, C. T. (1987). Child/ adolescent behavioral and emotional problems: lmplications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*, 213-222.

Achenbach, T.M. (1986). *Manual for the Teacher Report Form and teacher version of the Child Behavior Profile*. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1987). *Manual for the Youth Self-Report and Profile.* Burlington: University of Vermont, Department of Psychiatry.

Albrecht, G., Veerman, J. W., Damen, H., & Kroes, G. (2001). The Child Behavior Checklist for group care workers: A study regarding the factor structure. *Journal of Abnormal Child Psychology, 29*, 83-89.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arrindell, W. A., & Ettema, J. H. M. (1986). *SCL-90* (Dutch translation and adaptation). Lisse: Swets & Zeitlinger.

Barrett, P. M., Rapee, R. M., Dadds, M. M., & Ryan, S. M. (1996). Family enhancement of cognitive style in anxious and aggressive children. *Journal of Abnormal Child Psychology, 24*, 187-203.

Beidel, D. C., & Turner, S. T. (1997). At risk for anxiety: I. Psychopathology in the offspring of anxious parents. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 918-924.

Bidaut-Russell, M., Reich, W., Cottler, L., Robins, L.N., Compton, W.M., & Mattison, R.E. (1995). The Diagnostic Interview Schedule for Children (PC-DISC v.3.0): Parents and adolescents suggest reasons for expecting discrepant answers. *Journal of Abnormal Child Psychology, 23*, 641-659.

Biederman, J., Mick,E., & Faraone, S. V. (1998). Biased maternal reporting of child psychopathology? [Letter to the editor]. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 10-11.

Bird, H.R., Gould, M.S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child and Adolescent Psychiatry, 31,* 78-85.

Blackman, M. C., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology, 34,* 164-181.

Boyle, M. H., & Pickles, A. R. (1997). Influence of maternal depressive symptoms on ratings of childhood behavior. *Journal of Abnormal Child Psychology, 25,* 399-412.

Breslau, N., Davis, J. C., & Prabucky, K. (1988). Depressed mothers as informants in family history research: are they accurate? *Psychiatry Research, 24,* 345-359.

Briggs-Gowan, M. J., Carter, A. S., & Schwab-Stone, M. (1996). Discrepancies between mother, child and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology, 24,* 749-765.

Brink, L.T. ten (1998). *De ontwikkeling van kinderen tijdens een periode van klinische jeugdzorg* [Child development during a period of clinical treatment]. Doctoral dissertation, Vrije Universiteit, Amsterdam.

Campbell, L.F., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81-105.

Chilcoat, H. D., & Breslau, N. (1997). Does psychiatric history bias mothers' reports? An application of a new analytic approach. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 971-979.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.

Cohen, J., Cohen. P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioural sciences.* Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, P., Velez, N., Kohn, M., Schwab-Stone, M., & Johnson, J. (1987). Child psychiatric diagnosis by computer algorithm: Theoretical issues and empirical tests. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 631-638.

Conrad, M., & Hammen, C. (1989). Role of maternal depression in perceptions of child maladjustment. *Journal of Consulting and Clinical Psychology, 57,* 663-667.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and the Five Factor Inventory (NEO-FFI): Professional Manual.* Odessa, Florida: Psychological Assessment Resources Inc.

Costello, A.J., Edelbrock, C., Dulcan, M.K., Kalas, R., & klaric, S. (1987). *Diagnostic Interview Schedule for Children (DISC).* Pittsburgh: Western Psychiatric Institute and Clinic, School of Medicine, University of Pittsburgh

Darley, J.M., & Fazio, R.H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35,* 867-881.

Delsing, M. (2004). *Family justice and trust. Analysis of intergenerational, cross-sectional, and longitudinal family data by means of the social relations model.* Doctoral dissertation, Radboud University, Nijmegen.

Derogatis, L. R. (1977). *SCL-90: Administration, Scoring and Procedures Manual-I for the R(evised) Version.* Baltimore: John Hopkins University School of medicine.

Dix, T. (1991). The affective organization of parenting. Adaptive and maladaptive processes. *Psychological Bulletin, 110,* 3-25.

Dix, T., & Lochman, J.E. (1990). Social cognition and negative reactions to children: A comparison of mothers of aggressive and nonaggressive boys. *Journal of Social and Clinical Psychology, 9,* 418-438.

Dodge, K. A. (1986). A social information processing model of social competence in children. In M. Perlmutter (Ed.), *Minnesota Symposium on Child Psychology, Vol. 18* (pp. 77-125). Hillsdale: Erlbaum.

Dodge, K.A. (1993). Social cognitive mechanisms in the development of conduct disorders and depression. *Annual Review of Psychology, 44,* 559-584.

Dodge, K.A., Pettit, G.S., McClaskey, C.L., & Brown, M.M. (1986). Social Competence in Children. *Monographs of the Society for Research in Child Development, 51,* 1-85.

Downey, G., & Coyne, J. C. (1990). Children of depressed parents: An integrative overview. *Psychological Bulletin, 108,* 50-76.

Epkins, C. C. (1994). Peer ratings of depression, anxiety and aggression in inpatient and elementary school children: Rating biases and influence of rater's self-reported depression, anxiety and aggression. *Journal of Abnormal Child Psychology, 22,* 611-628.

Epkins, C. C., & Dedmon, A. M. M. (1999). An initial look at sibling reports on children's behavior: Comparisons with children's self-reports and relations with siblings' self-reports and sibling relationships. *Journal of Abnormal Child Psychology, 27,* 371-382.

Esmeijer, F.J., Veerman, J.W., & van Leeuwen, H.M.P. (1999). Gestructureerde interviews voor DSM-classificatie van psychische stoornissen bij kinderen en jeugdigen [Structured interviews for the DSM classification of psychological symptoms in children and adolescents]. *Tijdschrift voor Psychiatrie, 41,* 209-218.

Esmeijer, F.J., Veerman, J.W., ten Brink, L.T., & van Leeuwen, H.M.P. (1998). Gestructureerde interviews bij DSM-classificatie van probleemgedrag bij kinderen en jeugdigen [Structured interviews for the DSM classification of psychological symptoms in children and adolescents]. *Kind en Adolescent, 19,* 263-273.

Ferdinand, R. F., & Verhulst, F. C. (1998). *Gedragsvragenlijst voor kinderen (CBCL) – Direct Observatie Formulier* [Child Behavior Checklist (CBCL) – Direct Observation Form]. Rotterdam: Erasmus Universiteit.

Fergusson, D. M., Lynskey, M. T., & Horwood, L. J. (1993). The effect of maternal depression on maternal ratings of child behavior. *Journal of Abnormal Child Psychology, 21,* 245-268.

Fiske, D.W. (1987). Construct invalidity comes from method effects. *Educational and Psychological measurement, 47,* 285-336.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652-670.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception.* San Diego: Academic Press.

Funder, D.C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science, 2,* 31-39.

Funder, D.C., & Colvin, C.R. (1991). Explorations in behavioral consistency: Properties of persons, situations and behaviors. *Journal of Personality and Social Psychology, 60,* 773-794.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Goodman, S. H., & Brumley, H. E. (1990). Schizophrenic and depressed mothers: Relational deficits in parenting. *Developmental Psychology, 26,* 31-39.

Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review 106,* 458-490.

Graham, P., & Rutter, M. (1986). The reliability and validity of psychiatric assessment of the child. II. Interview with the parent. *British Journal of Psychiatry, 114,* 581-592.

Hall, J.A. (1990). *Nonverbal sex differences: Accuracy of communication and expressive style.* Baltimore: John Hopkins University Press.

Hammen, C., Burge, D., & Stansbury, K. (1990). Relationship of mother and child variables to child outcomes in a high-risk sample: A causal modeling analysis. *Developmental Psychology, 26,* 24-30.

Herjanic, B., & Reich, W. (1982). Development of a structured psychiatric interview for children: Agreement between child and parent on individual symptoms. *Journal of Abnormal Child Psychology, 10,* 307-324.

Hinshaw, S.P., Han, S.S., Erhardt, D., & Huber, A. (1992). Internalizing and externalizing behavior problems in preschool children: Correspondence among parent and teacher ratings and behavior observations. *Journal of Clinical Child Psychology, 21,* 143-150.

Hodges, K., McKnew, D., Cytryn, L., Stern, L., & Kline, J. (1982). The Child Assessment Schedule (CAS) diagnostic interview: a report on reliability and validity. *Journal of the American Academy of Child and Adolescent Psychiatry, 21,* 468-473.

Hoekstra, H. A., Ormel, J., & De Fruyt, F. (1996). *NEO-PI-R en NEO-FFI Persoonlijkheidsvragenljsten: Handleiding* [NEO-PI-R en NEO-FFI Personality Inventories: Manual]. Lisse: Swets & Zeitlinger.

Hofman, A.L., & Tates, J.G. (19887). *Informant variance. Een exploratief follow-up onderzoek naar verschillen in beoordelingen van ouders en kinderen, van gedragsproblemen en gezinsfunctioneren* [Informant variance. An explorative follow-up study on the differences in the evaluation by parents and children, of behavioral problems and family dysfunction]. Unpublished master's thesis, Vrije Universiteit, vakgroep Kinder- en Jeugdpsychologie, Amsterdam.

Ingersoll, B. D., & Eist, H. I. (1998). Are depressed mothers biased reporters? [Letter to the editor]. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 681.

Jansen, M.G., & Oud, J.H.L. (1993). *Residentiële hulpverlening geëvalueerd. Eindrapport* [Residential care evaluated. Final report]. Nijmegen: Katholieke Universiteit, Instituut voor Orthopedagogiek.

Jensen, P. S., Traylor, J., Xenakis, S. N., & Davis, H. (1988). Child psychopathology rating scales and interrater agreement: I. Parents' gender and psychiatric symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 442-450.

Johnston, J., & Short, K. H. (1993). Depressive symptoms and perceptions of child behavior. *Journal of Social and Clinical Psychology, 12,* 164-181.

Kashani, J. H., Vaidya, A. F., Soltys, S. M., Dandoy, A. C., Katz, L. M., & Reid, J. C. (1990). Correlates of anxiety in psychiatrically hospitalized children and their parents. *American Journal of Psychiatry, 147,* 319-323.

Kazdin, A. E., Esveldt-Dawson, K., Unis, A., & Rancurello, M. (1983). Child and parent evaluation of depression and aggression in psychiatric inpatient children. *Journal of Abnormal Child Psychology, 11,* 401-413.

Kazdin, A.E. (1988). The diagnosis of childhood disorders: assessment issues and strategies. *Behavioral Assessment, 10,* 67-94.

Kazdin, A.E., & Bass, D. (1988). Parent, teacher and hospital staff evaluations of severely disturbed children. *American Journal of Orthopsychiatry, 58,* 512-523.

Kendziora, K.T., & O'Leary, S. G. (1998). Appraisals of child behavior by mothers of problem and nonproblem toddlers. *Journal of Abnormal Child Psychology, 26,* 247-255.

Kenny, D.A. (1994*). Interpersonal perception.* New York: Guildford Press.

Kenny, D.A., & LaVoie, L. (1984). The social relations model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 141-179). New York: Academic Press.

Kloosterman, M.V., & Veerman, J.W. (1996). *CBCL - TRF onderzoek Boddaertcentra 1996. Gegevens over opname en vertrek* [CBCL – TRF research Boddaertcentra 1996. Data on admission and discharge]. Utrecht: VOG.

Kloosterman, M.V., & Veerman, J.W. (1997). *CBCL - TRF onderzoek Boddaertcentra 1997. Gegevens over opname en vertrek* [CBCL – TRF research Boddaertcentra 1997. Data on admission and discharge]. Utrecht: VOG.

Kloosterman, M.V., & Veerman, J.W. (1999). *Boddaert belicht. Eindrapportage interventie-evaluatieonderzoek Boddaert* [Boddaert elucidated. Final report Boddaert intervention-evaluation research]. Utrecht: VOG.

Kolko, D. J., & Kazdin, A. E. (1993). Emotional/behavioral problems in clinic and nonclinic children: Correspondence among child, parent and teacher reports. *Journal of Child Psychology and Psychiatry, 34,* 991-1006.

Kroes, G., & van der Doef, P.LM. (1994). Beoordeling van gedragsproblemen met de CBCL/TRF. Populatiekenmerken en veranderingen tijdens de behandeling [Evaluation of behavioral problems using the CBCL/TRF. Population characteristics and changes during treatment]. In H.M. Pijnenburg, C.M. van Rijswijk & J.W. Veerman (Eds.), *Pedologisch Jaarboek 1994* (pp. 73-87). Delft: Eburon.

Kroes, G., Veerman, J. W., & De Bruyn, E. E. J. (2000). Realiteit en vertekening bij het beoordelen van probleemgedrag van kinderen [Reality and bias in the evaluation of child problem behavior]. In J. D. Bosch, R. J. van der Gaag, A. J. J. M. Ruijssenaars, & A. Vyt (Eds.), *Jaarboek Ontwikkelingspsychologie, Orthopedagogiek en Kinderspychiatrie 4 (2000-2001)* (pp. 87-118). Houten/Diegem: Bohn Stafleu Van Loghum.

Kroes, G., Veerman, J. W., & De Bruyn, E. E. J. (2003). Bias in parental reports? Maternal psychopathology and the reporting of problem behavior in clinic-referred children. *European Journal of Psychological Assessment, 19,* pp. 196–204.

Kroes, G., Veerman, J. W., & De Bruyn, E. E. J. (2005). The impact of the Big Five personality traits on reports of child behavior problems by different informants. *Journal of Abnormal Child Psychology, 33,* 231-240.

Kroes, G., Veerman, J. W., & De Bruyn, E. E. J. (2005). *The Perception of Child Behavior Problems: The Role of Acquaintanceship, Informant Personality, and Context.* Manuscript submitted for publication.

Levesque, M.J., & Kenny, D.A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *Journal of Personality and Social Psychology, 65,* 1178-1187.

Loeber, R., Green, S., Lahey, B.B., & Stouthamer-Loeber, M. (1989). Optimal informants on childhood disruptive behaviors. *Development and Psychopathology, 1*, 317-337.

Lorber, M. F., O'Leary, S. G., & Kendziora, K. T. (2003). Mothers' overreactive discipline and their encoding and appraisals of toddler behavior. *Journal of Abnormal Child Psychology, 31*, 485-494.

McCombs Thomas, A., Forehand, R., Armistead, L., Wierson, M., & Fauber, R. (1990). Cross-informant consistency in externalizing and internalizing problems in early Adolescence. *Journal of Psychopathology and Behavioral Assessment, 12,* 255-262.

McCrae, R. R., & Costa, P. T. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research.* (pp. 139-153). New York: Guilford Press.

Meijer, R.W.J., & Veerman, J.W. (1989). Gedragsbeoordeling door verschillende informanten. Literatuurbespreking en eigen onderzoeksgegevens [Behavioral evaluation by different informants]. *Kind en Adolescent, 10*, 183-194.

Mesman Schultz, K. (1978). *De COM-procedure. Een methode om voor de kinderbeschermingspupillen een advies te bepalen* [The COM-procedure. A method for reaching an advice for child protection youth]. Amsterdam: Stichting Interuniversitair Instituut voor Sociaal-Wetenschappelijk Onderzoek/SISWO.

Mick, E., Santangelo, S. L., Wypij, D., & Biederman, J. (2000). Impact of maternal depression on ratings of comorbid depression in adolescents with attention deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 314-319.

Mokros, H.B., Poznanski, E., Grossman, J.A., & Freeman, L.N. (1987). A comparison of child and parent ratings of depression for normal and clinically referred children. *Journal Child Psychology and Psychiatry, 28,* 613-627.

Moretti, M. M., Fine,S., Haley, G., & Marriage, K. (1985). Childhood and adolescent depression: Child-report versus parent-report information. *Journal of the American Academy of Child and Adolescent Psychiatry, 24,* 298-302.

Muntinga, I. (1999). *Het construeren en beoordelen van video-opnames van kindgedrag in een klinische setting: een inleidende studie* [The construction and rating of child behavior videotapes in a clinical setting: A preliminary study]. Unpublished master's thesis, Radboud University, Nijmegen.

Najman, J. M., Williams, G. M., Nikles, J., Spence, S., Bor, W., O'Callaghan, M., LeBroque, R., & Andersen, M. (2000). Mothers' mental illness and child behavior problems: Cause-effect association or observational bias? *Journal of the American Academy of Child and Adolescent Psychiatry, 39,* 592-602.

Offord, D.R., Boyle, M.H., Racine, Y., Szatmari, P., Fleming, J.E, Santford, M., & Lipman, E.L. (1996). lntegrating assessment data from multiple informants. *Joumal of the American Academy of Child and Adolescent Psychiatry, 35*, 1078-1085.

Orobio de Castro, B., & Bosch, J.D. (1998). Sociale-informatieverwerking door jongens met gedragsproblemen [Social information processing in boys with behavior problems]. In J.D. Bosch e.a. (Eds.), *Jaarboek Ontwikkelingspsychologie, orthopedagogiek en kinderpsychiatrie3, 1998-1999* (pp. 212-241). Houten: Bohn Stafleu Van Loghum.

Patterson, G.R. (1982). *Coercive family process.* Eugene: Castalia.

Patterson, G.R. (1986). Maternal rejection: Determinant or product for deviant child behavior? In W. Hartup & Z. Rubin (Eds.), *Relationships and Development* (pp. 73-92). Hillsdale: Erlbaum.

Peters, M. (1998). *Bias in de beoordeling van probleemgedrag bij kinderen door moeders met psychische problematiek* [Bias in the evaluation of child problem behavior by mothers with psychological symptoms]. Unpublished master's thesis, Katholieke Universiteit, vakgroep Orthopedagogiek, Nijmegen.

Phares, V. (1997). Accuracy of informants: Do parents think that mother knows best? *Journal of Abnormal Child Psychology, 25,* 165-171.

Phares, V., Compas, B.E., & Howell, D.C. (1989). Perspectives on child behavior problems: Comparison of children's self-reports with parent and teacher reports. *Psychological Assessment, 1*, 68-71.

Piacentini, J.C., Cohen, P., & Cohen, J. (1992). Combining discrepant diagnostic information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology, 20,* 51-63.

Power, T.J., Andrews, T.J., Eiraldi, R.B., Doherty, B.J., Ikeda, M.J., DuPaul, G.J., Landau, S., (1998). Evaluating Attention Deficit Hyperactivity Disorder using multiple informants: The incrementalutility of combining teacher with parent reports. *Psychological Assessment, 10,* 250-260.

Reich, W., & Welner, Z. (1988). *Revised version of the Diagnostic Interview for Children and Adolescents (DICA-R).* St Louis: Department of Psychiatry, Washington University School of Medicine.

Richters, J. E (1992). Depressed mothers as informants about their children. A critical review of the evidence for distortion. *Psychological Bulletin, 112,* 485-499.

Richters, J. E., & Pellegrini, D. (1989). Depressed mothers' judgments about their children: An examination of the depression-distortion hypothesis. *Child Development, 60,* 1068-1075.

Rosnow, R., & Rosenthal, R. (1997). *People studying people: Artifacts and ethics in behavioral research.* New York: Bedford, Freeman, & Worth.

Rutter, M. (1990). Commentary: Some focus and process considerations regarding effects of parental depression on children. *Developmental Psychology, 26,* 60-67.

Sawyer, M. G., Streiner, D. L., & Baghurst, P. (1998). The influence of distress on mothers' and fathers' reports of childhood emotional and behavioral problems. *Journal of Abnormal Child Psychology, 26,* 407-414.

Saylor, C., Finch, A., Baskin, C., Furey, W., & Kelly, M. (1984). Construct validity for measures of childhood depression: Application of multitrait-multimethod methodology. *Journal of Consulting and Clinical Psychology, 52,* 977-985.

Silverman, W.K., & Eisen, A.R. (1992). Age differences in the reliability of parent and child reports of child anxious symptomatology using a structured interview. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 717-724.

Slotboom, A., Elphick, E., & Mervielde, I. (2000). Temperaments- en persoonlijkheidsonderzoek bij kinderen: de belangrijkste kenmerken in kaart gebracht [Research on temperament and personality: An overview of the most important characteristics]. In J. D.Bosch R. J. van der Gaag, A. J. J. M. Ruijssenaars, & A. Vyt (Eds.), *Jaarboek Ontwikkelingspsychologie, Orthopedagogiek en Kinderspychiatrie 4 (2000-2001)* (pp. 59-86). Houten/Diegem: Bohn Stafleu Van Loghum.

Snarr, J. D., Strassberg, Z., & Slep, A. M. S. (2003). Making faces: Testing the relation between child behavior problems and mothers' interpretations of child emotion expressions. *Journal of Abnormal Child Psychology, 31,* 371-380.

Sourander, A., & Piha, J. (1997). Parent, teacher, and clinical ratings on admission to child psychiatric inpatient treatment. A study of cross-informant correlations using the CBCL, Rutter scales, and the CGAS. *Nordic Journal of Psychiatry, 51,* 365-370.

Stanger, C., & Lewis, M. (1993). Agreement among parents, teachers, and children on internalizing and externalizing problems. *Journal of Clinical Child Psychology, 22,* 107-115.

Stavrakaki, C., Vargo, B., Roberts, N., & Boodoosingh, L. (1987). Concordance among sources of information for ratings of anxiety and depression in children. *Journal of the American Academy of Child and Adolescent Psychiatry , 26,* 733-737.

Stevens, J. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Strassberg, Z. (1995). Social information processing in compliance situations by mothers of behavior-problem boys. *Child Development, 66,* 376-389.

Strassberg, Z. (1997). Levels of analysis in cognitive base of maternal disciplinary dysfunction. *Journal of Abnormal Child Psychology, 25,* 209-215.

Treiber, F.A., & Mabe, P.A. (1987). Child and parent perceptions of childrens' psychopathology in psychiatric outpatient children. *Journal of Abnormal Child Psychology, 15,* 115-124.

Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent-child relationships? *Journal of Abnormal Child Psychology, 31,* 13-28.

Valla, J.-P., Bergeron, L., Breton, J.-J., Gaudet, N., & Berthiaume, C. (1993). Informants, correlates, and child disorders in a clinical population. *Canadian Journal of Psychiatry, 38,* 406-411.

Van Westerlaak, J. M., Kropman, J. A., & Collaris, J. W. M. (1975). *Beroepen-klapper* [Occupational Index]. Nijmegen: Instituut voor Toegepaste Sociologie.

Verhulst, F. C., Ende, J. van der, & Koot, H. M. (1996). *Handleiding voor de CBCL/4-18* (Manual for the CBCL/4-18). Rotterdam: Sophia Kinderziekenhuis, Afdeling Kinder- en Jeugdpsychiatrie.

Verhulst, F. C., Ende, J. van der, & Koot, H. M. (1997). *Handleiding voor de Teacher's Report Form (TRF)* [Manual for the Teacher's Report Form (TRF)]. Rotterdam: Erasmus Universiteit/Sophia Kinderziekenhuis, Afdeling Kinder- en Jeugdpsychiatrie.

Verhulst, F.C., & Akkerhuis, G.W. (1989). Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4-12. *Journal of Child Psychology and Psychiatry, 30,* 123-136.,

Verhulst, F.C., & van der Ende, J. (1991). Assessment of child psychopathology: relationships between different methods, different informants and clinical judgment of severity. *Acta Psychiatrica Scandinavica, 84*, 155-159.

Verhulst, F.C., Althaus, M., & Berden, G.F.M.G. (1987). The Child Assessment Schedule: parent-child agreement and validity measures. *Journal of Child Psychology and Psychiatry, 28*, 455-466..

Weis, R., & Lovejoy, M. C. (2002). Information processing in everyday life: Emotion-congruent bias in mothers' reports of parent-child interactions. *Journal of Personality and Social Psychology, 83,* 216-230.

Youngstrom, E., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology, 67*, 905-916.

Youngstrom, E., Loeber, R., & Stouthamer-Loeber, M. (2000). Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology, 68*, 1038-1050.

Samenvatting

DE WAARNEMING VAN PROBLEEMGEDRAG BIJ KINDEREN

De rol van de persoonlijkheid van de informant en de context

Deze dissertatie was gericht op de vraag wie de meest nauwkeurige informatie over het gedrag van een kind kan geven, onder welke omstandigheden (zie Hoofdstuk 1). In de klinische praktijk wordt informatie over het functioneren van een kind gewoonlijk verschaft door verschillende informanten, zoals de ouders, leerkrachten, andere professionals, en de kinderen zelf. Over het algemeen wordt er een lage overeenstemming gevonden tussen de verschillende beoordelingen van het gedrag van een kind. Eén mogelijke verklaring voor deze lage overeenstemming tussen informanten is gelegen in het feit dat verschillende typen informanten (bijv. ouders, vrienden, leerkrachten) het gedrag van een kind vaak onder zeer verschillende omstandigheden waarnemen. Met andere woorden, gedragsproblemen bij kinderen kunnen situatiespecifiek zijn. Een tweede mogelijke verklaring voor de lage overeenstemming tussen verschillende informanten kan gelegen zijn in de eigenschappen van de informanten zelf. Met andere woorden, de informatie over gedragsproblemen bij kinderen kan ook informantspecifiek zijn.

In Hoofdstuk 2, getiteld *Realiteit en vertekening bij het beoordelen van probleemgedrag van kinderen*, werd een overzicht gegeven van de stand van zaken met betrekking tot onderzoek naar de waarneming van gedragsproblemen. Er is aanzienlijke discussie over de mogelijke invloed van psychopathologie bij ouders, en in het bijzonder depressie bij moeders, op de beoordeling van kindgedrag. Zo zou depressie bij moeders volgens diverse onderzoekers leiden tot een vertekend oordeel over de gedragsproblemen van hun kinderen, maar volgens anderen is dit vooralsnog onvoldoende aangetoond. Bovendien is het mogelijk dat ook andere persoonlijkheidseigenschappen een vertekenende invloed kunnen hebben op de oordelen van ouders over kindgedrag, hoewel men het er algemeen over eens is dat het bewijs voor een dergelijke vertekening zwak is. Een groot probleem bij de beoordeling

van gedragsproblemen vormt het gebrek aan criteria om de juistheid en validiteit van oordelen over kindgedrag aan af te meten. Aangezien iedere informant het gedrag van een kind vanuit zijn of haar eigen perspectief bekijkt, is het verzamelen van zo veel mogelijk informatie over het gedrag van een kind nog de beste manier om hiermee om te gaan. En hoewel er verschillende procedures zijn ontwikkeld om informatie uit verschillende bronnen te combineren en te structureren, is geen van deze procedures tot nu toe de beste gebleken.

Om het beoordelen van gedragsproblemen bij kinderen te verbeteren hebben we het voorstel gedaan om alle factoren die van invloed lijken te zijn op de waarneming en beoordeling van kindgedrag onder te brengen in één enkel, meer algemeen, sociaal waarnemingsmodel — het *Realistic Accuracy Model* (Funder, 1995). Het *Realistic Accuracy Model* omvat factoren die zowel betrekking hebben op de situatie als op de informant, en bovendien verenigt dit model informatie uit verschillende domeinen van de sociale psychologie en persoonlijkheidsleer. We hebben beargumenteerd dat dit model ook voor onderzoek naar het beoordelen van kindgedrag een relevant en vruchtbaar theoretisch raamwerk biedt.

In Hoofdstuk 3, getiteld *Vertekening in ouderlijke rapportages? Psychopathologie bij moeders en het rapporteren van probleemgedrag van opgenomen kinderen,* presenteerden we de eerste van drie empirische studies die zijn opgezet om de nauwkeurigheid en vertekening van de waarneming van probleemgedrag bij kinderen te onderzoeken. We onderzochten het verband tussen verschillende vormen van psychische klachten bij moeders en de rapportage door moeders van internaliserende en externaliserende gedragsproblemen bij hun kinderen bij een onderzoeksgroep van 68 jongens in de leeftijd van 6 tot 12 jaar, met behulp van de Klachtenlijst (SCL-90), de ouderversie van de Gedragsvragenlijst voor kinderen (CBCL), en de leerkrachtversie van de Gedragsvragenlijst voor kinderen (TRF). Aangezien er geen absolute standaard bestaat voor het identificeren van gedragsproblemen, beschouwden we de rapportages van leerkrachten en groepsleiders over de gedragsproblemen bij de kinderen als de meest nauwkeurige bron van informatie, en deze rapportages namen we als criterium om te vergelijken met de rapportages van de moeders. Als we zouden vinden dat de rapportages van de moeders afwijken van die van de leerkrachten en groepsleiders, en als we bovendien een systematisch verband zouden vinden tussen deze afwijkingen en de mate van depressie bij moeders, dan zou aannemelijk zijn dat er sprake is van vertekening in de waarneming van gedragsproblemen door moeders op grond van hun depressie.

Met behulp van multipele regressie-analyses vonden we substantiële verbanden tussen verschillende vormen van psychische klachten bij moeders en de rapportages door moeders van internaliserende gedragsproblemen bij hun kinderen, na deze gecontroleerd te hebben voor de criteriumbeoordelingen. Met andere woorden, moeders met meer psychische klachten hadden de neiging om meer internaliserende gedragsproblemen bij hun kinderen te rapporteren dan moeders met minder psychische klachten, ook wanneer rekening werd gehouden met de mate van internaliserend probleemgedrag die gerapporteerd werd door de

leerkrachten en groepsleiders. Dit wijst op een vertekening in de waarneming van internaliserende gedragsproblemen door moeders met meer dan gemiddelde psychische klachten. Wat betreft het rapporteren van externaliserend probleemgedrag werden slechts geringe of niet-significante correlaties gevonden. Het feit dat er vooral bij het rapporteren van internaliserende gedragsproblemen vertekening optrad, is in overeenstemming met de sociale attributietheorie, die voorspelt dat er meer vertekening optreedt bij het observeren van meer ambigue stimuli.

In ons eerste onderzoek werden beoordelingen van gedragsproblemen door verschillende informanten als criterium gebruikt om te vergelijken met de beoordeling door de moeders. Echter, de verschillende typen informanten observeerden en beoordeelden het gedrag van de kinderen onder heel verschillende omstandigheden (d.w.z., thuis, in de klas, of in de leefgroep). Daardoor kunnen de afwijkingen in de beoordelingen tussen de verschillende typen informanten weliswaar worden toegeschreven aan informantkenmerken zoals psychische klachten bij de moeders, maar net zo goed kunnen deze afwijkingen worden toegeschreven aan situationele factoren. Dat wil zeggen, het is heel goed mogelijk dat kinderen van depressieve of angstige moeders thuis *feitelijk* meer gedragsproblemen laten zien dan in de klas of in de leefgroep. In dat geval zou er geen sprake zijn van een vertekende waarneming van gedragsproblemen door depressieve of angstige moeders, maar juist van een accurate waarneming van een grotere mate van kindproblemen thuis — mogelijk onder invloed van depressie of angst bij de moeders —, in vergelijking met school of de kliniek. Om te controleren voor de verschillen in beoordelingen die veroorzaakt worden door situationele factoren is het noodzakelijk om alle informanten exact hetzelfde kindgedrag te laten beoordelen. Anders gezegd, het is noodzakelijk om het te beoordelen kindgedrag te standaardiseren.

In Hoofdstuk 4, getiteld *De invloed van de Big Five persoonlijkheidsdimensies op de rapportage van probleemgedrag bij kinderen door verschillende informanten*, presenteerden we een tweede onderzoek, waarbij we gebruik hebben gemaakt van video-opnamen van het gedrag van 55 kinderen in de leeftijd van 6 tot 12 jaar. Deze video-opnamen, die gemaakt werden in leefgroepen in de kliniek, dienden als een gestandaardiseerde steekproef van kindgedragingen om door de verschillende informanten te worden beoordeeld. De gedragsbeoordelingen door de moeders, leerkrachten en groepsleiders werden met elkaar vergeleken. Bovendien onderzochten we de invloed van persoonlijkheidskenmerken van de moeders, leerkrachten en groepsleiders — gemeten met de NEO-FFI, een vragenlijst die de vijf belangrijkste dimensies van de persoonlijkheid, te weten neuroticisme, extraversie, openheid, altruïsme en consciëntieusheid, meet — op hun beoordeling van gedragsproblemen bij de kinderen. Alle informanten bekeken en beoordeelden dus dezelfde 17 minuten durende video-opname van het gedrag van een hun bekend kind. Bovendien beoordeelden getrainde observatoren, die het kind niet kenden, de video-opnames om te zorgen voor een onafhankelijk criterium bij de beoordeling van de gedragsproblemen.

De resultaten van ons tweede onderzoek lieten zien dat, a) de moeders over het algemeen meer problemen rapporteerden dan de leerkrachten of de groepsleiders, b) de informanten die het kind kenden over het algemeen meer problemen rapporteerden dan de informanten die de kinderen niet kenden (d.w.z., de onafhankelijke observatoren), en c) een hogere mate van neuroticisme bij de informanten samenhing met het rapporteren van meer gedragsproblemen in het geval van de leerkrachten en groepsleiders, maar niet in het geval van de moeders. Bovendien rapporteerden meer extraverte en meer open groepsleiders over het algemeen *minder* gedragsproblemen dan groepsleiders met een gemiddelde mate van extraversie en openheid. Tenslotte werd er geen verband gevonden tussen de mate van altruïsme of consciëntieusheid en het rapporteren van meer of minder gedragsproblemen. De resultaten van het tweede onderzoek suggereren dus dat professionals die met kinderen werken niet immuun zijn voor vertekening bij hun waarneming van gedragsproblemen.

De invloed van bekendheid met het kind op het rapporteren van kindgedrag werd verder onderzocht in onze derde en laatste studie, beschreven in Hoofdstuk 5 onder de titel *De waarneming van gedragsproblemen bij kinderen: de rol van bekendheid, persoonlijkheid van de informant, en context.* We hebben gebruik gemaakt van Funder's (1995) theorie over het beoordelen van persoonskenmerken om de rol van bekendheid te analyseren, en om een aantal methodologische problemen die we in eerder onderzoek naar vertekende waarneming bij informanten tegenkwamen opnieuw onder de loep te nemen. Bovendien onderzochten we de invloed van verschillende situationele contexten — d.w.z., thuis in het geval van de moeders, de leefgroep in het geval van de groepsleiders, en video-opnamen gemaakt in de leefgroep in het geval van beide typen informanten — op de waarneming van kindgedrag en mogelijke vertekening daarvan. Moeders en groepsleiders beoordeelden video-opnames van zowel een bekend als een onbekend kind, en tevens het gedrag van het bekende kind in de dagelijks leven thuis (in het geval van de moeders) of in de leefgroep (in het geval van de groepsleiders). De videobanden werden ook beoordeeld door onafhankelijke beoordelaars, die geen van de kinderen kenden.

In overeenstemming met de theorie over de invloed van bekendheid op het beoordelen van persoonskenmerken (*Acquaintanceship effect*; Funder, 1999), bleek dat moeders meer gedragsproblemen waarnamen dan onafhankelijke beoordelaars, maar dit was alleen het geval bij de beoordeling van bekende kinderen en niet bij de beoordeling van onbekende kinderen. Groepsleiders daarentegen rapporteerden meer gedragsproblemen dan onafhankelijke beoordelaars bij zowel bekende als onbekende kinderen. Bovendien bleek dat neuroticisme bij moeders verband hield met hun beoordeling van gedragsproblemen bij hun eigen kinderen, maar dit was alleen het geval bij de beoordeling van het gedrag van hun kind in de context van het dagelijks leven, en niet bij de beoordeling van (de video-opname van) het gedrag van hun kind in de context van de leefgroep. Daarentegen bleek de mate van neuroticisme bij groepsleiders verband te houden met hun beoordeling van gedragsproblemen bij de kinderen

in beide contexten (d.w.z., zowel bij de beoordeling van het gedrag in de dagelijkse omgang als bij de beoordeling van de video-opname van het kind).

In Hoofdstuk 6, getiteld *Samenvatting en Discussie*, hebben we de resultaten op een rij gezet en een aantal discussiepunten besproken. Het belangrijkste methodologische probleem bij onderzoek naar vertekening in de waarneming van gedragsproblemen is het ontbreken van een 'gouden standaard' om de aanwezigheid van een gedragsprobleem aan af te meten. Ook onafhankelijke, getrainde beoordelaars die vaak als criteriuminformanten gebruikt worden, kunnen onnauwkeurig zijn in hun beoordeling vanwege onbekende oorzaken en in onbekende mate, volgens Richters (1992). In ons onderzoek hebben we aanvankelijk gebruik gemaakt van leerkrachten en groepsleiders, en later van getrainde beoordelaars om criteriumbeoordelingen te leveren. Hoewel het gebruik van getrainde beoordelaars, in combinatie met een gestandaardiseerde observaties van kindgedrag, zeker een vooruitgang betekent vanwege het controleren van situationele factoren, is het vanuit het oogpunt van het *Acquaintanceship effect* (Funder, 1999) nog maar de vraag of het oordeel van getrainde informanten die het kind niet kennen wel accurater is dan dat van informanten die het kind wel van nabij kennen.

In de meeste eerder gerapporteerde studies die gebruik maakten van onafhankelijke beoordelaars en video-opnames van kindgedrag bleken onafhankelijke beoordelaars over het algemeen veel minder gedragsproblemen te rapporteren dan ander informanten die het kind wel kenden. In ons laatste onderzoek bleek hetzelfde het geval te zijn; bovendien bleken de moeders in ons onderzoek bij hun eigen kinderen *wel*, maar bij onbekende kinderen *niet* meer gedragsproblemen te onderkennen dan onafhankelijke beoordelaars. Tezamen suggereren deze resultaten dat bekenden wellicht inderdaad betere en nauwkeuriger beoordelaars zijn dan onbekenden, ook al zijn die onbekenden getraind in het observeren en beoordelen van kindgedrag. Dat betekent dat Richters' methodologische eisen voor onderzoek naar vertekening in de waarneming van gedragsproblemen — namelijk, dat rapportages van onafhankelijke beoordelaars als criteriumbeoordeling dienen te worden gebruikt — wellicht dienen te worden aangevuld met de eis dat ook informanten die criteriumbeoordelingen leveren bekend zijn met het te beoordelen kind.

Een theoretisch model dat mogelijkerwijs een zinvolle aanvulling zou kunnen bieden op het door ons gehanteerde *Realistic Accuracy Model* van Funder (1995) is het *Social Relations Model* (Kenny & Lavoie, 1984). Dit model biedt namelijk een methode om de oordelen van gezinsleden over elkaar te analyseren, wat deze methode bij uitstek geschikt maakt om vragen te beantwoorden als: a) In welke mate heeft een gezinslid de neiging om alle medegezinsleden op een bepaalde, algemene manier te beoordelen? b) In hoeverre komt het oordeel van een bepaald gezinslid over een ander overeen met het oordeel van de andere gezinsleden? c) In hoeverre is het oordeel van een bepaald gezinslid over een ander gezinslid uniek binnen het gezin? (Vgl. Delsing, 2004). Een dergelijke analysemethode vergt weliswaar de nodige investering van de gezinsleden omdat een complete set van beoordelingen van alle

gezinsleden over en weer vereist is, maar voldoet beter dan eerder onderzoek naar vertekende waarneming aan de eis dat alle informanten bekend zijn met het te beoordelen kind. En een dergelijke analysemethode biedt tevens de mogelijkheid om alle informanten exact hetzelfde kindgedrag in dezelfde context te laten beoordelen, namelijk thuis, om uiteindelijk aan te tonen of de persoonlijkheid van moeders (en mogelijk ook andere gezinsleden) nu wel of niet een vertekenend effect heeft op hun beoordeling van gedragsproblemen bij een kind.

Ons onderzoek naar vertekening in de waarneming van gedragsproblemen door moeders onder invloed van hun eigen persoonlijke problematiek begon met het vergelijken van de oordelen van moeders, leerkrachten en groepsleiding. We dachten aan te kunnen tonen dat moeders met psychische klachten inderdaad meer gedragsproblemen waarnemen bij hun kinderen dan er in *feitelijk* zijn (lees: dan er door anderen worden waargenomen). We dachten ook dat de oordelen van leerkrachten en groepsleiders, die we als criterium hanteerden, nauwkeuriger zouden zijn dan die van de moeders en niet onderhevig aan mogelijke invloeden van hun eigen persoonlijkheid. Bij nader onderzoek bleek geen van beide uitgangsstellingen houdbaar: moeders bleken, althans bij het beoordelen van video-opnames van hun kind, niet beïnvloed te worden door hun eigen persoonlijkheid, terwijl dat juist wel het geval bleek bij leerkrachten en groepsleiders. Met name de persoonlijkheidsdimensies zelfverzekerdheid (of haar tegenhanger, neuroticisme) en extraversie bleken van invloed op de beoordeling van kindgedrag. Dit komt overeen met eerder onderzoek naar de eigenschappen van goede beoordelaars (Funder, 1999).

Tenslotte willen we opmerken dat ons onderzoek geen definitieve antwoorden kan geven op de oorspronkelijke vraag van dit dissertatieonderzoek, welke informant de meest nauwkeurige en valide beoordeling geeft van gedragsproblemen van een kind, onder welke omstandigheden. Daarvoor is het beslist nodig om onze conclusies te staven met onderzoek bij grotere populaties, met andere criteria voor gedragsproblemen, en met onderzoeksdesigns die met name de gedragsproblemen thuis en de mogelijke invloed van ouderlijke problematiek daarop nauwkeuriger onder de loep nemen. Bovendien bleken zowel de beoordeling van gedragsproblemen als de studie daarvan complexer dan we hadden gedacht. Wat we wel hebben aangetoond is dat zowel informantkenmerken, waaronder een aantal specifieke persoonlijkheidsfactoren, als ook de situationele context waarin gedrag wordt waargenomen van grote invloed kunnen zijn op het rapporteren van gedragsproblemen bij kinderen. Daarmee is, om in termen van Funder te spreken, de studie van nauwkeurigheid of vertekening bij de beoordeling van gedragsproblemen complexer geworden, maar ook realistischer. Achenbach et al. concludeerden al in 1987 dat bij het beoordelen van gedragsproblemen van kinderen bij voorkeur informatie uit verschillende bronnen verzameld dient te worden, om gedrag van het kind in verschillende situaties en vanuit verschillende invalshoeken te belichten. Wij zouden daaraan toe willen voegen dat bij het verzamelen van die informatie rekening gehouden dient te worden met de persoonlijkheid van de verschillende informanten en met de context waarin het gedrag is waargenomen.

Dankwoord
(Acknowledgements)

Dit proefschrift zou nooit tot stand zijn gekomen zonder de royale steun en medewerking van Stichting De Waarden. Ik dank dan ook in de eerste plaats directie en medewerkers van De Waarden hartelijk voor hun collegiale, praktische, en morele ondersteuning bij de uitvoering van mijn promotieonderzoek. Alleen dankzij hun hulp was het mogelijk om de gegevens van ouders te verzamelen, video-opnamen te maken in de leefgroepen en op school, en de logistiek voor de interviews met een groot aantal deelnemers rond te krijgen. Ook ben ik het bestuur van de *Stichting tot Dienstverlening aan De Waarden* zeer erkentelijk voor hun financiële bijdrage, die het mij mogelijk maakte gedurende een aantal jaren een groot deel van mijn tijd als behandelaar voor dit onderzoek vrij te maken. En tenslotte wil ik Frans Kaufman nog danken, die indertijd als directeur van De Waarden aan de wieg heeft gestaan van dit promotieproject.

Diverse mensen hebben een belangrijke rol gespeeld bij de totstandkoming van dit proefschrift door geheel belangeloos aan het onderzoek deel te nemen. Hen wil ik in het bijzonder voor hun bijdrage bedanken. Daarbij denk ik in de eerste plaats aan de kinderen die bereid waren zich te laten filmen, en hun ouders die niet alleen bereid waren om vragenlijsten over hun kinderen in te vullen en video-opnamen te beoordelen, maar ook om informatie over zichzelf te verstrekken. Ik dank hen hartelijk voor de tijd die zij in het onderzoek hebben willen investeren en voor hun vertrouwen — in de verwachting dat zij daarmee op termijn wellicht het belang van een nauwkeuriger diagnostiek bij kinderen met probleemgedrag zouden dienen. Ik hoop dat ik die verwachting niet heb beschaamd en dat zij zich in de uitkomsten van dit onderzoek kunnen vinden.

Ook de groepsleiders en leerkrachten die bereid waren videobanden te beoordelen en daarnaast informatie te verstrekken over hun eigen persoonlijkheid hebben een belangrijke bijdrage geleverd aan dit onderzoek. Tot nu toe waren er geen gegevens bekend over de invloed van de persoonlijkheid van de professionals zelf op hun beoordeling van kindgedrag. Het feit dat de groepsleiders en leerkrachten van De Waarden bereid waren deze informatie te leveren betekende dan ook een belangrijke stap voorwaarts in dit soort onderzoek. Ook hen dank ik hartelijk voor hun medewerking en het door hen geschonken vertrouwen.

Wat betreft de uitvoering van het onderzoek ben ik veel dank verschuldigd aan een aantal studenten dat in het kader van hun stage en scriptie een belangrijke bijdrage hebben geleverd bij het inzamelen van vragenlijsten, bedenken van een videoscript, maken van opnames, interviewen van ouders, groepsleiders en leerkrachten, verzamelen van gegevens,

berekenen en analyseren van uitkomsten, en het verlenen van diverse hand- en spandiensten. Brechtje van Gent, Barbara Gutteling, Isabel Jansen, Froukje Kamstra, Ingrid Muntinga, Marjon Peters, en Hanneke Stokkermans – hartelijk dank voor jullie onmisbare, enthousiaste, en zeer gewaardeerde inzet. Ook dank ik de onafhankelijke beoordelaars voor hun inzet bij de training en uiteindelijke beoordeling van de video-opnamen.

Bij het analyseren van de gegevens heb ik een aantal keren dankbaar gebruik gemaakt van de methodologische kennis van John van den Bercken, die elegante oplossingen aandroeg voor statistische vraagstukken. Ik waardeer de enthousiaste en voortvarende wijze waarop hij mijn vragen onder handen nam en dank hem hartelijk daarvoor. Verder ben ik Lee Ann Weeks erkentelijk voor haar bijdrage aan het schrijven van dit proefschrift. Zij heeft de Engelse tekst gecorrigeerd, fraaie zinsconstructies aangedragen, en waardevolle adviezen gegeven voor het verbeteren van de leesbaarheid en logica van de tekst.

Tenslotte, lest best, wil ik mijn beide promotoren danken voor hun jarenlange begeleiding bij het opzetten, uitvoeren en publiceren van dit onderzoek. Jan Willem en Eric, ik prijs me gelukkig met jullie deskundige coaching en grote betrokkenheid bij dit onderzoek, en heb deze als een voortdurende stimulans ervaren. Ik denk dan ook met plezier terug aan de vele begeleidingssessies waarin we de voortgang besproken (en soms heftig bediscussieerd) hebben. Ik heb er veel van geleerd, met name waar het de kunst van het ordenen, structureren, en weglaten betreft. Ook bij het schrijven van de artikelen voor dit proefschrift had ik mij geen betere begeleiders kunnen wensen. Hartelijk dank daarvoor!

Curriculum vitae

Gert Kroes was born in Genemuiden, The Netherlands, in 1952. He attended the Johannes Calvijn Lyceum in Kampen and then went on to study sociology and philosophy at the Vrije Universiteit in Amsterdam in 1971, psychology at the Universiteit van Amsterdam in 1972, and to graduate with a degree in psychology from the Katholieke Universiteit Nijmegen in 1979.

After obligatory military service, Gert was a research assistant in the Vakgroep Klinische Psychologie at the Katholieke Universiteit Nijmegen where he participated in comparative research on the effects of client-centered and behavior therapy. At the same time, he undertook postgraduate on-the-job training to become a psychotherapist at the RIAGG (Mental Health Institute) in Nijmegen. From 1982 to 1986, he was employed as a psychotherapist by Huize Alexandra, a juvenile penitentiary in Almelo.

From 1986 to 2002, Gert worked as a developmental psychologist at the Paedologisch Instituut, Nijmegen, a department of the Stichting de Waarden, which is a multidisciplinary center for child care, special education, professional training, and child-care-related research. During this time, he participated in research on such topics as the assessment of child behavior problems, the differential diagnosis of pervasive developmental disorders, and the outcomes of social skills training. He also authored or co-authored various publications and presentations on these topics. In 1998, he embarked on his PhD project examining variation in the assessment of child behavior problems under the auspices of the Stichting de Waarden in Nijmegen. He is currently a researcher at Praktikon, another department of the Stichting de Waarden located within the Academisch Centrum Sociale Wetenschappen of the Radboud Universiteit Nijmegen (the Academic Center for Social Sciences at the Radboud University Nijmegen).